

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:36:54 ; Search time 175 Seconds
(without alignments)
2651.561 Million cell updates/sec

Title: US-10-619-992-2

Perfect score: 6215

Sequence: 1 MPVQAQWTEFLSCPTCYNE.....MSDEKNDFLKPVANGRMVNS 1191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6215	100.0	1191	9	US-09-921-099-2
2	6207	99.9	1191	9	US-09-921-099-4
3	611	9.8	522	15	US-10-104-047-3665
4	490	7.9	95	9	US-09-867-550-1346
5	311	5.0	377	9	US-09-864-761-48290
6	207	3.3	2377	18	US-10-840-512-196
7	203	3.3	1267	15	US-10-092-900A-128
8	200.5	3.2	728	17	US-10-732-923-6679
9	198.5	3.2	3586	15	US-10-334-143-77
10	196.5	3.2	1248	16	US-10-437-963-126940
11	195.5	3.1	123	9	US-09-864-761-42017

12	195	3.1	2135	15	US-10-288-798-9	Sequence 9, Appli
13	195	3.1	2135	15	US-10-362-892-9	Sequence 9, Appli
14	195	3.1	2382	14	US-10-196-935A-2	Sequence 2, Appli
15	195	3.1	2382	15	US-10-052-648A-40	Sequence 40, Appl
16	195	3.1	2382	15	US-10-336-472-230	Sequence 230, App
17	195	3.1	2382	16	US-10-408-765A-1404	Sequence 1404, Ap
18	192.5	3.1	1390	15	US-10-092-900A-224	Sequence 224, App
19	191.5	3.1	2816	15	US-10-240-145-145	Sequence 4, Appli
20	191	3.1	1129	18	US-10-840-512-119	Sequence 119, App
21	191	3.1	4262	17	US-10-704-781-4	Sequence 3, Appli
22	191	3.1	4493	17	US-10-704-781-3	Sequence 1940, Ap
23	190.5	3.1	842	15	US-10-369-493-1940	Sequence 125036,
24	190	3.1	922	16	US-10-437-963-125036	Sequence 3473, Ap
25	190	3.1	1247	14	US-10-128-714-3473	Sequence 1085, Ap
26	185.5	3.0	1343	16	US-10-408-765A-1085	Sequence 2, Appli
27	185.5	3.0	2325	16	US-10-663-433-2	Sequence 425, App
28	182.5	2.9	4019	9	US-09-738-973-425	Sequence 425, App
29	182.5	2.9	4019	9	US-09-854-133-425	Sequence 425, App
30	182.5	2.9	4019	14	US-10-144-649A-425	Sequence 62, Appl
31	182	2.9	2766	11	US-09-964-956-62	Sequence 589, App
32	181.5	2.9	2210	16	US-10-719-993-589	Sequence 7758, Ap
33	181	2.9	1130	14	US-10-032-585-7758	Sequence 156135,
34	180.5	2.9	1405	16	US-10-437-963-156135	Sequence 104121,
35	180.5	2.9	1609	16	US-10-437-963-104121	Sequence 140753,
36	180	2.9	409	16	US-10-437-963-140753	Sequence 160683,
37	180	2.9	1424	16	US-10-437-963-160683	Sequence 14, Appl
38	179	2.9	2783	9	US-09-816-669A-14	Sequence 2228, Ap
39	179	2.9	2783	16	US-10-846-648-14	Sequence 32198, A
40	179	2.9	2789	16	US-10-408-765A-2228	Sequence 181703,
41	178.5	2.9	800	14	US-10-029-386-32198	Sequence 188268,
42	178	2.9	419	16	US-10-437-963-181703	Sequence 195106,
43	177.5	2.9	457	16	US-10-437-963-188268	Sequence 2246, Ap
44	177.5	2.9	606	16	US-10-437-963-195106	
45	177.5	2.9	723	15	US-10-104-047-2246	

ALIGNMENTS

RESULT 1
US-09-921-099-2
; Sequence 2, Application US/09921099
; Patent No. US20020107372A1
; GENERAL INFORMATION:
; APPLICANT: Hefeneider, Steven
; APPLICANT: Merkins, Louise
; APPLICANT: Bennett, Robert
; APPLICANT: Seiss, Donald
; TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene
; FILE REFERENCE: 00-617-A
; CURRENT APPLICATION NUMBER: US/09/921,099
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-099-2

Query Match	100.0%	Score 6215;	DB 9;	Length 1191;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1191;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	1	MPVQAQWTEFLSCPICYNEFDENVHKPISLGCSTVCKTCLNKLHRKACPPDQTAINTD	60	
Db	1	MPVQAQWTEFLSCPICYNEFDENVHKPISLGCSTVCKTCLNKLHRKACPPDQTAINTD	60	
Qy	61	IDVLPVNFALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVEDLALYKPLSGGKGVAS	120	
Db	61	IDVLPVNFALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVEDLALYKPLSGGKGVAS	120	
Qy	121	LNQSALSRPMQRKLVTLVNCQLVEEGRVRAARSLGERTVTTELQHNQNPQQLSANL	180	

Db 121 LQNSALSRPWQRKLVTLVNCQVEEGRVRAARAASLGERTVTELLIHOHQNPQQLSANL 180
QY 181 WAAVARGCOFLGPANQBEALKLVLALDGSALSRKVLVLFVQVLEPRFPQASKTSIG 240
Db 181 WAAVARGCOFLGPANQBEALKLVLALDGSALSRKVLVLFVQVLEPRFPQASKTSIG 240
QY 241 HVVQLLYRASCCKVTKRDEDSIMOLKEEFRSYEARLREHDAQIVHIAEAGLRISPEQM 300
Db 241 HVVQLLYRASCCKVTKRDEDSIMOLKEEFRSYEARLREHDAQIVHIAEAGLRISPEQM 300
QY 301 SLLYGDIAHAKSHMQSIIDKLQSPESFAKSVQELTIIVLQRTGDPANLRLRPHLELLANI 360
Db 301 SLLYGDIAHAKSHMQSIIDKLQSPESFAKSVQELTIIVLQRTGDPANLRLRPHLELLANI 360
QY 361 DNPDAVSPTWQLENAWAVKTVVHGLVDFIQNTYSRKGHETPQOPNSKYKTSMCRDLR 420
Db 361 DNPDAVSPTWQLENAWAVKTVVHGLVDFIQNTYSRKGHETPQOPNSKYKTSMCRDLR 420
QY 421 OQGGCPRGTCNCTFAHSELEKYLRLNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480
Db 421 OQGGCPRGTCNCTFAHSELEKYLRLNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480
QY 481 GSTETTGKIVPSTNGISNAENSVQLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540
Db 481 GSTETTGKIVPSTNGISNAENSVQLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540
QY 541 SVTENKIGSPKTPVSNVAATSAGPSNVGTSLNSVPQKSPFLTRVPVYPHSENIQYFQ 600
Db 541 SVTENKIGSPKTPVSNVAATSAGPSNVGTSLNSVPQKSPFLTRVPVYPHSENIQYFQ 600
QY 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPESLPPASMPYADHYST 660
Db 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPESLPPASMPYADHYST 660
QY 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRRIWRPPMYQRDDIIRNSLPPM 720
Db 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRRIWRPPMYQRDDIIRNSLPPM 720
QY 721 DVNHSVYQTSLRERYNSLDGYYSVACQPPSPRTTVPLPRBPCGHLKTSCEQIRRKPD 780
Db 721 DVNHSVYQTSLRERYNSLDGYYSVACQPPSPRTTVPLPRBPCGHLKTSCEQIRRKPD 780
QY 781 QNAQYHTQKAPLSVSTLPVATQSPPLSFVDFRADFSESVSGTKFEEDHLSHYSPWS 840
Db 781 QNAQYHTQKAPLSVSTLPVATQSPPLSFVDFRADFSESVSGTKFEEDHLSHYSPWS 840
QY 841 CGTIGSCINAIDSEPKDVIANNAVLMDLSDGDKRRVHLFETQRTKEEDPIIPFSDGP 900
Db 841 CGTIGSCINAIDSEPKDVIANNAVLMDLSDGDKRRVHLFETQRTKEEDPIIPFSDGP 900
QY 901 IISKWGAISRSGRTGYHTTDPVQATASQGSATKPISVSDYVPVNAVDSRWSSYNGEATS 960
Db 901 IISKWGAISRSGRTGYHTTDPVQATASQGSATKPISVSDYVPVNAVDSRWSSYNGEATS 960
QY 961 SAHYVERDRFIYTDLSCHBKSHSTGDLISLEQQAQKNSLILQREANALAMQOKNSLDE 1020
Db 961 SAHYVERDRFIYTDLSCHBKSHSTGDLISLEQQAQKNSLILQREANALAMQOKNSLDE 1020
QY 1021 GRHLTLNLKSKETELRANGELQSDYTDATDTKPDRIEELSALDTPDQSGSEPIEEL 1080
Db 1021 GRHLTLNLKSKETELRANGELQSDYTDATDTKPDRIEELSALDTPDQSGSEPIEEL 1080
QY 1081 DIQLGTSQNDQLLNGMAVENGHVPQOQKEPPKQKQSLGEDHVLILEQKTIIPVTSCEF 1140
Db 1081 DIQLGTSQNDQLLNGMAVENGHVPQOQKEPPKQKQSLGEDHVLILEQKTIIPVTSCEF 1140
QY 1141 SQPLPVSISNASCLPITTSVAGNLIJLKTVMSEDKNDFLKPVANGKMYNS 1191
Db 1141 SQPLPVSISNASCLPITTSVAGNLIJLKTVMSEDKNDFLKPVANGKMYNS 1191

US-09-921-099-4
; Sequence 4, Application US/09921099
; Patent No. US20020107372A1
; GENERAL INFORMATION:
; APPLICANT: Hefeneider, Steven
; APPLICANT: Merkins, Louise
; APPLICANT: Bennett, Robert
; APPLICANT: Seigs, Donald
; TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene
; FILE REFERENCE: 00-617-A
; CURRENT APPLICATION NUMBER: US/09/921,099
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-099-4

Query Match 99.9%; Score 6207; DB 9; Length 1191;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPVQAAQWTEFLSCPICYNEDFNHVKPISLGCSTHTVCKTCLNKLHRRKACPDQTAINTD 60
Db 1 MAVQAAQWTEFLSCPICYNEDFNHVKPISLGCSTHTVCKTCLNKLHRRKACPDQTAINTD 60
QY 61 IDVLVNFALLQLVGAQVPDQHSIKLSNLGENKHVEYAKKVEDLALYLKPLSGGKGVAS 120
Db 61 IDVLVNFALLQLVGAQVPDQHSIKLSNLGENKHVEYAKKVEDLALYLKPLSGGKGVAS 120
QY 121 LQNSALSRPWQRKLVTLVNCQVEEGRVRAARAASLGERTVTELLIHOHQNPQQLSANL 180
Db 121 LQNSALSRPWQRKLVTLVNCQVEEGRVRAARAASLGERTVTELLIHOHQNPQQLSANL 180
QY 181 WAAVARGCOFLGPANQBEALKLVLALDGSALSRKVLVLFVQVLEPRFPQASKTSIG 240
Db 181 WAAVARGCOFLGPANQBEALKLVLALDGSALSRKVLVLFVQVLEPRFPQASKTSIG 240
QY 241 HVVQLLYRASCCKVTKRDEDSIMOLKEEFRSYEARLREHDAQIVHIAEAGLRISPEQM 300
Db 241 HVVQLLYRASCCKVTKRDEDSIMOLKEEFRSYEARLREHDAQIVHIAEAGLRISPEQM 300
QY 301 SLLYGDIAHAKSHMQSIIDKLQSPESFAKSVQELTIIVLQRTGDPANLRLRPHLELLANI 360
Db 301 SLLYGDIAHAKSHMQSIIDKLQSPESFAKSVQELTIIVLQRTGDPANLRLRPHLELLANI 360
QY 361 DNPDAVSPTWQLENAWAVKTVVHGLVDFIQNTYSRKGHETPQOPNSKYKTSMCRDLR 420
Db 361 DNPDAVSPTWQLENAWAVKTVVHGLVDFIQNTYSRKGHETPQOPNSKYKTSMCRDLR 420
QY 421 OQGGCPRGTCNCTFAHSELEKYLRLNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480
Db 421 OQGGCPRGTCNCTFAHSELEKYLRLNKKINATVTRFPLLNKVGNNVTVTAGNVSIV 480
QY 481 GSTETTGKIVPSTNGISNAENSVQLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540
Db 481 GSTETTGKIVPSTNGISNAENSVQLISRSTDSLRALETVKVKGVGANGQNAAGPSAD 540
QY 541 SVTENKIGSPKTPVSNVAATSAGPSNVGTSLNSVPQKSPFLTRVPVYPHSENIQYFQ 600
Db 541 SVTENKIGSPKTPVSNVAATSAGPSNVGTSLNSVPQKSPFLTRVPVYPHSENIQYFQ 600
QY 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPESLPPASMPYADHYST 660
Db 601 DRTQIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPESLPPASMPYADHYST 660
QY 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRRIWRPPMYQRDDIIRNSLPPM 720
Db 661 FSPDRMNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRRIWRPPMYQRDDIIRNSLPPM 720
QY 721 DVNHSVYQTSLRERYNSLDGYYSVACQPPSPRTTVPLPRBPCGHLKTSCEQIRRKPD 780

```
Db 721 DVWSSVYQTSURYNLDGYYSVACQPPSPRTVPLRPFCGHLKTSCEQIRRKPD 780
Qy 781 QWAOYHTOKAPLVSSTLPVATOSPPTPPSPPLFSVDPRADFSVSGTKFEEDHLSHYSPWS 840
Db 781 QWAOYHTOKAPLVSSTLPVATOSPPTPPSPPLFSVDPRADFSVSGTKFEEDHLSHYSPWS 840
Qy 841 CQTIGSCINAIIDSEPKDVIANAVLMDLSDGDKVRRVHLFTQRTKEEDPIIFSDGP 900
Db 841 CQTIGSCINAIIDSEPKDVIANAVLMDLSDGDKVRRVHLFTQRTKEEDPIIFSDGP 900
Qy 901 IISKWGAISRSRTGHTTDPVQATASQGSATKPIISVDYVYVNAVDSRWSSSYNEATS 960
Db 901 IISKWGAISRSRTGHTTDPVQATASQGSATKPIISVDYVYVNAVDSRWSSSYNEATS 960
Qy 961 SAHYVERDRFIVTDLSGHKKHSTGDLLELQQAQKSNLLQREANALAMQOKNSLDE 1020
Db 961 SAHYVERDRFIVTDLSGHKKHSTGDLLELQQAQKSNLLQREANALAMQOKNSLDE 1020
Qy 1021 GRHLTLNLSKEIERNGLQSDYTEDATDTKPDRIELESALOTDEPDGQSEPIEIL 1080
Db 1021 GRHLTLNLSKEIERNGLQSDYTEDATDTKPDRIELESALOTDEPDGQSEPIEIL 1080
Qy 1081 DIQLGISSQNDQLLGMVANGHPVQHQKPPKQKQSLGSDHVLBEQKTLPTVSCF 1140
Db 1081 DIQLGISSQNDQLLGMVANGHPVQHQKPPKQKQSLGSDHVLBEQKTLPTVSCF 1140
Qy 1141 SLPVSVISNASCLPITTSVSAGNLILKTHVMSDKNDFLKPVANGKMVNS 1191
Db 1141 SLPVSVISNASCLPITTSVSAGNLILKTHVMSDKNDFLKPVANGKMVNS 1191

RESULT 3
US-10-104-047-3665
; Sequence 3665, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3665
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3665

Query Match 9.8%; Score 611; DB 15; Length 522;
Best Local Similarity 34.1%; Pred. No. 2e-33;
Matches 187; Conservative 76; Mismatches 187; Indels 98; Gaps 27;

Qy 618 YTPPPPTVAGVACVPRFRVRSNNVSESLPPASMEYADHYSTFSDRDMNSPIQPPPP 677
Db 2 YTTTTPP-----QCVSRRFVR-----PPPSAPEAPPYLDHYPPYL-OERVNSQYGTQPO 49
Qy 678 QPYGVPVPPVPSGMVAPVYDSRRIRWRPPMYORDDIIRSNSLPPMDVMHSSV--YOTSRLR 735
Db 50 Q-----YPTPI-----YPSHYDGRVTPAPSYTRESIFRESPI-PIEIPPAVPSVPSRER 100
Qy 736 YNSLDGYYSVACQP-----PS-----EPRTTVPLPREPCGHLKTSCEQIRRKPDQWQYHT 787
Db 101 YQIESYYPVAPHTQIRPSYLREPVSRLPPPPQP-----HPSLDELHRRRKEIMAQLEB 156
Qy 788 QKAPLVSSTLPVATQSTPPSPPLFSVDPRADFSVSGTKFEEDHLSHYSPWSCTGIGSC 847
Db 157 RK---VISPPPPFA-PSPTLP-PTFHPPEEFLDEDLKVG-KYKGNDSYQSPWNSCDTIGSY 210
Qy 848 INAIIDSEPKDVIANAVLMDLSDGDKVRRVHLFTQRTKE--EDPIIFSDGPIISKW 905
```

```
Db 211 IGTDAKPKDVAAAGSVEMMNVESKGM--RQRLDLQRAAETSDDDLIPGEDRTVSRF 268
Qy 906 GAISRSRTGHTTDPVQATASQGSATKPIISVDYVYVNAVDSRW--SSYG--NEATSS 961
Db 269 GAISRTSKTIIYQAGPQMAAPQAGPTKSIINISDYSY--GTHGGMGASPYSPHQNIPSQ 326
Qy 962 AHYVERDRFIVTDLSGHKK--HSSTGDLLELQ-----AKNSLLQOR 1004
Db 327 GHFSERERISMSEVASHGKPLPSAEREQLRLELQQLNHQISQOTQLRGPEAVSNRLVLQR 386
Qy 1005 EANALAMQO-----KWSLDEGRHLTLNL--LSKEIERNGLQSDYTEDATDTKPD 1054
Db 387 EANTLAGOSQPPPPPPKPGKMISEQLSLELHQVEREIGKTRLSME--NCSLDMK-- 443
Qy 1055 RDIELELSALDTDE--PDGQSEPIEILDIQLGISSQNDQLLGMVANGHPVQHQKPEP 1112
Db 444 -----SKLNTSKQAENGQPEPQNKV-----PAEDLTITFSDVPNGSALTQENISL 488
Qy 1113 PKQKQSL 1120
Db 489 LSNKTSSL 496

RESULT 4
US-09-867-550-1346
; Sequence 1346, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: NO. US20020082206A1el Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1346
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-550-1346

Query Match 7.9%; Score 490; DB 9; Length 95;
Best Local Similarity 100.0%; Pred. No. 3.5e-26;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1097 MAVENGHPVQHQKPPKQKQSLGSDHVLBEQKTLPTVTCFSPQPLPVISNASCLPI 1156
Db 1 MAVENGHPVQHQKPPKQKQSLGSDHVLBEQKTLPTVTCFSPQPLPVISNASCLPI 60
Qy 1157 TTSVSGNLIILKTHVMSDKNDFLKPVANGKMVNS 1191
Db 61 TTSVSGNLIILKTHVMSDKNDFLKPVANGKMVNS 95

RESULT 5
US-09-864-761-48290
; Sequence 48290, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48290
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL136170.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: SWISSPROT HIT: Q01970, EVALUE 2.60e+00
; OTHER INFORMATION: EST_HUMAN HIT: AW962574.1, EVALUE 4.00e-32
US-09-864-761-48290
```

```
Query Match 5.0%; Score 311; DB 9; Length 67;
Best Local Similarity 89.6%; Pred.No. 5.4e-14;
Matches 60; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 254 VTKRDEDSLMQKEEFREYALRREHDAQIVHIAEAGIRISPEQWSSLLYCDLAHKS 313
DB 1 VTKRDEDSLMQKEEFREYALRREHDSQIVQIAEAGLRAPDQWSSLLYGDOSHKSH 60
QY 314 MQSIIDK 320
DB 61 MQSIIDK 67
```

RESULT 6

```
US-10-840-512-196
; Sequence 196, Application US/10840512
; Publication No. US20050125852A1
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARYDCZAK, GLEN
```

```
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840,512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 196
; LENGTH: 2377
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-840-512-196

Query Match 3.3%; Score 207; DB 18; Length 2377;
Best Local Similarity 19.3%; Pred.No. 0.00022;
Matches 216; Conservative 152; Mismatches 424; Indels 326; Gaps 46;

QY 48 KACPFDQTAINT-----DIDLVPVNFALLQLVGAQVDPHQSIKLSNLGEN 92
DB 375 KRASPAKSVIGTPEFMAPEMYEEKYDESVDYVAFGCMCLEMATSEYP----- 421
QY 93 KHYEVAKKCVEDLALYLKPLSGKGVAASLNOSAL-----SRPWOKULVLVN 139
DB 422 -----YSECQNAAQIYRRVTSVK-PASFDFKVAIPEVKEIIIEGCIRQNKDERYSIKDLIN 475
QY 140 QLVVEEGRVAMRAARSIGERTVTTELILQHNQPOOLSANLWAAVRARGCQFLGPAMQEE 199
DB 476 HAFQOETGVRVELAEDDGEKIAIKLWRIEDIKKLK-----GKYKONE 520
QY 200 ALKLVLLALEDGSAISRKVLVLFVVRLEPRFPQASKTSIGHVVO-----LLYRAS 250
DB 521 AIEF-----SFLERDVPEDVAGEMVSGVCEGDHKTAKAIDRV 563
QY 251 CFKVTKRDEDSLMQKEEFREYALRREHDAQIVHIAEAGLRISPEQWSSLLYCDLAH 310
DB 564 LIK-RKREQROLVREEQEKQKESSFKQNEQASVS-QAGI-----QLSAASTGIPTA 617
QY 311 KSHMQSIIDKQSPESFAKSQVELTIVLQRTCDPANLRPLHLLANIDNPDAVSPT 370
DB 618 PATSASVSTQVEPEPEADQHQ-----LOYQQPSSISVL-S-DGTIDSGQGS 662
QY 371 WEQLENAMVAVKTVVHGLVDFITQNZSRKGHET-----PQOPNSKYKTSMCRD 418
DB 663 SVFTESRVSSQQTVSYG-----SQHEQAHSTGTAPGHTVSSIOAQSQPHGVYPPSSMAQ 716
QY 419 LRQOQGCPRGNTCTFAHSQEELEKYRLNKKINATVTFPLIN--KVGNNNTVTTAGNV 476
DB 717 GQNQOQ-PSSSLAGVLSQ-----PIHQPOQOQGIQPT----- 747
QY 477 ISVIGSTETTKIVPSTNGISNAENSVSOLISRSTDSTLRALETYKVKVKGANGQNAAG 536
DB 748 -----VPSQAV---QVSLPQAASS-----EG 767
QY 537 PSADSVTENKIGSPKTPVSNVAATSAGPSN--VGTELNSVPQKSPFLTRVPVPPHSE 594
DB 768 TTAQPVSPQVSAGTQLPVSTVAIVQGEPHIPVSTQPSVVPVHGAHFLPMGQPIPTSL 827
QY 595 NIQYQDPRTQIPFEVPQY--PQTGYPPPTPVAGVAPCVPRFVRNNVPSSLLPPASM 652
DB 828 LPQY---PVSQIPISTPHVSTAQTGFSSVPITMAAGINQ--PLLTASSATASSIPGGSP 882
QY 653 PYADHYST-FSPDRDMNSSPY-----QPPPPQPYGVP-----PVPSC--MYAPVD 696
DB 883 VVPNQLPTLLQPVNQLSQVHPQLLQPTTVOSIG-IPANLGOAEGPLPSGGVLYQG-PP 940
QY 697 SRIWRPMPYQDDIIRNSLPPMDVMHSSVY-QTSLRERYNSLDGYYSVACQPPSEPT 755
DB 941 SR---LPQYPCDSNIAPSSNVASVCIHSTVLAPPSPMTEALATQGYPTTVQPVESPT 997
QY 756 TVPLPREPCGHLKTSCEBOIRRKPDQWQAYHTQKAPLVSS-----TL 797
```

Db 998 LVPV-GSVGGQVVS-----QPAVSLTQOPPTTSSQAVLESTQGVSOAAPPETQ 1046
QY 798 PVATQPTPSPPLFS-----VDFRADFS-----SVSGTKPEEDHLGSHYSWSCGTIG 845
Db 1047 PITQSQPTQVPLVTSADSHSDVAGMSDGNENAPSSGRHGRRTTKHYRK---SVRS 1103
QY 846 SCINAIDSEPKOVIAN-----SNAVLMDDSGDVKRRVHLFETQRTKEEDPIIPFSDGP 900
Db 1104 RSRHEKTSRPKILNVKNGDRVVEQLETHNRKMVTFKFDLDGDNPEIATIMVND 1163
QY 901 IISKWGAISRS-----RTGYHTTDPVQ-----ATASQ 928
Db 1164 IL-----AIERESVAQVREIEKADMLSEDVSEVPEGGQLESQKDDYGPFGSKULE 1219
QY 929 GSATKPIISDYVPYVNVDSRWSSYGNEATSAHYVDRDRFIVTDLGHRKHS---TG 985
Db 1220 GEFKQPIAVSSMQQIGVPTS-----SLTQVHSAGR-RFIVSPVESRLRESKVFTS 1271
QY 986 DLLS-LELOQAKNSLLLOREANALAMQKWNLSDEGR 1022
Db 1272 DISDPVAVASTQAPGNLHSHSASSLSLQQAFAFSELKHGQ 1309

RESULT 7

US-10-092-900A-128
; Sequence 128, Application US/10092900A
; Publication No. US20040043382A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Gorman, Linda
; APPLICANT: Miller, Charles E.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Liu, Yi
; APPLICANT: Anderson, David W.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Catterton, Elina
; APPLICANT: Leite, Mario W.
; APPLICANT: Zhong, Haihong
; APPLICANT: Aleobrook, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: No. US20040043382A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-290C
; CURRENT APPLICATION NUMBER: US/10/092,900A
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: USSN 60/274,322
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/283,675
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: USSN 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: USSN 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: USSN 60/274,191
; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: USSN 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: USSN 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: USSN 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: USSN 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: USSN 60/287,424
; PRIOR FILING DATE: 2001-04-30
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 768
; SEQ ID NO 128
; LENGTH: 1267
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-900A-128

Query Match 3.3%; Score 203; DB 15; Length 1267;
Best Local Similarity 19.8%; Pred. No. 0.00016;
Matches 190; Conservative 116; Mismatches 337; Indels 318; Gaps 46;

QY 423 GCGPRG---TNCCTFAHSQEELEK-----YRLRNKKIN----- 451
Db 21 GRCPRNSLYSDCIIIEKTVLQKQNEGFGFVLRGAKADTPIEEFTPTTAPFALQKLESV 80
QY 452 -----ATVRTFLLNKGVNNTVTTTGNVISVI--GSTETTGKIVPSTNGI----- 496
Db 81 DEGGVAVQAGLRTGDFLIEVNNENVVKGVRQVNMIRQGNHVLKVVTVTLENLPDDT 140
QY 497 -----SNAENSVSQILSRSTDTSLRALETVKVKGKVGANGQNAAGPSADSVTENKI 547
Db 141 ARKKAPPPKRAPTTALTLSKSMTSSELELDKPEI-----VPASKPS---RAAENMA 191
QY 548 GSPPKTPVSNVAATSAGPSN-----VCTELNSVPQKSSPLTRVPVYPVPHSENLOYQDPR 603
Db 192 VEP-----RVATIKQRPSSRCFFPAGSDMNVSGLTGP-RGRGPTVPPRLSGSQSVYE-R 243
QY 604 TQIPFEVPOYPOTGYVPPPP---TVPAG-----VAPCVPRFVRSNNVPE 644
Db 244 QGIAVMTPTVGS---PKAPFLIGIPGTWRROKSIGITEERQFLAPPMLKFTFSLMPD 300
QY 645 SS--LPPASMPYADHYSTFSPRDMNSSPYQPPPPQ---YG-PVPVPS--GMVAPVYDS 697
Db 301 TSEDIPPPP-----QSVPPSPPPSPPTTNCPSKPTPRVYGTIKPAFNQ 344
QY 698 RRIWRPPMYQRDDIIRSNLPLPMDVMHSSVYQTSLEARYNSLDG--YYSVACOPPSEPRT 755
Db 345 NSAAKVSPATRSDDTVATM-----MREKGMVFRRELDRY-SLDSSELYSRNAGFOANFRN 397
QY 756 -TVPLPREPCGHLKTSCEQI-----RRKPDQWAQYHTQKAPLVSSSTLPVAT----- 801
Db 398 KRQMPENPYSEVGKIASKAVVYPAKARKGMLVKQSNVEDSPSEKTSIPITIIIVKEP 457
QY 802 -----QSPTPPSPFLSVDFRADFSSEVSGTKFEEDHLGSHYSWSCGT 843
Db 458 STSSSGKSSQSGSMEIDPOAPEPPSQL-----RPDESLSVTS-----SPFAAAI 500
QY 844 IGSCINAIDSEPKOVIANNSA-VLMDLDSGD-----VKRVLHLPETQRTKSED-- 891
Db 501 AG-----AVRDREKRELRNNSPAFLSTDLGDEVDVGLGPPAPTRPRSMFEEGFADESDA 556
QY 892 -----PIIPFSDGPIISKWGAISR-----SSRTGYHTTDPVQATASQSA-- 931
Db 557 EQLSSPMPSATPREPNHFVGAEASAPGEAGPLNSTSKAQGPSSPAVPSASSGTAGP 616
QY 932 -----TXPISV-----SDYVP-----YVNV-----AVD 948
Db 617 GNVVHPLTGRLLDPPSSPLALALSARDAMKESQGGPKGEAPKADLNKPLVYIDTKMRPSLD 676
QY 949 SRWSSYGNBAT--SSAHYVERDRPIVTDLSGHRKHSTGDL--SLELOQAKNSLLLORE 1005
Db 677 AGFPTVTRQNTGRPLRRQETENKYEYTDLGRDRKGGDKKNMLIDIMDTSOOKSAGLLMVT 736

Db 1689 SLLQPKGPFPAKQWGGADTLSTQTMGKSCNGLNVEERTVNLPAKPGTGVKIVIDRSQ 1748
QY 363 NPDAVSPTWEOLENA-----WVAVKTVVHGLVDFIQY-SRKGHETPQOPNSK 410
Db 1749 NDALSKSLRRLSNHHKSNVLESGDKTGWDTTGEQHLRDEYMSRQLTSTSVKQNL 1808
QY 411 YKTS--MCRDLR-----QQGCGPRGNCTF--AHSQEELEKVLRLNKKINATVRTP 458
Db 1809 TRKSDRAVELKDDVFNISQAGTKVGKQOTVELRNDHQWEGFHKSPKTKNLIK--- 1865
QY 459 LLNKGVNNTVTTAGNVLSIGSTETTKIIPSTNGISNAENSVLSRSTDSLTRAL 518
Db 1866 -----ILTDTOSS-KPSPTOHPVSMVPGTGYDLSDGFKQKLLQ 1904
QY 519 ETV---KKVGKVGANGN-----AAGPSADSVTENKIG----- 548
Db 1905 ETYKSNKDKKKININLOPMWQLLPVEQDTSNVTMKVSEKSHNTFKATNKKRETDVHLKS 1964
QY 549 -----SPPKTPVSNVAATSAGSNVGTLSNVPSQSPFLTRVVPVPHSENIQFQDP 602
Db 1965 QDFLMKTNSTGLKMAERSLNBNFNPENNVKSECEPLPPSPPPPPPS-----NA 2016
QY 603 RTQIPFEVQYQOTGYPPPTPVAGVAPCVPRFVS-----NNVPSSLPASM 652
Db 2017 SSBIEFPLP-----PPPLM---MFPEKNGFLPSLSTEKIKAEFEPGLPLPP--- 2062
QY 653 PYADHYSTSPDRMNSPPVOPPOPYGVPVPSGMVAPVYDSRRIRWPPMPYQDDII 712
Db 2063 PVDEKS-----ERSSSMLPPPP-----PTPS-----QXPAHL 2093
QY 713 RNSLPPMDVMHSS--VYQTSLRERYNSLDGYYSV-----ACQPPSEPTTVPLPREPC 764
Db 2094 LSSSAPE--KHSGDFMQYVSQKEASNSQSAKIITKGTGLVLP-----TLPKPLP- 2144
QY 765 GHUK-----TSCEBQIRKPDQ-----WAQVHTO-KAPLVSSITLPAV 801
Db 2145 KHIKDNKNDFSPKVELATLSLDMCKITTSKDKQVMVMTSSEHTETKQNVISKSLDERK 2204
QY 802 QSTPSPPLFSVDFRADFSFSGTKFEEDH-----LSHSPWSGCTIGSCINAIIDSEP 855
Db 2205 Q-----LSIDSANCLSHTVPGTSAPRKQIAPLIKSHSPFESSG-----QQNP 2247
QY 856 KOVIANSNAVIMDLSDGVKRRVHLFTQRTKKEEDPIIPFSDGPIISKWGAISRSTRG 915
Db 2248 KPYMRKFTPLM---TABEKYRQKKEIEKQKQ-----SSYNIKVTQSQN 2291
QY 916 YHTD-----PVQATASQGSATKPISVSDVVPVYVNAVDSRWSSYGN-----AT 959
Db 2292 QHTEVEKEMLQKT-----NEEVSLSGIDSECTVVQSPSGQSNARILGVCSNDQLST 2345
QY 960 SSAHYVERDFIVTDLGHRKHSSTGDL--SLELOQAKSNLLO--REANALAMQOK- 1014
Db 2346 TSPETVAAGKHLHVAASEDKMKXKEVLQSSDIMQSKACEIKOSHOBECSTQOQOKK 2405
QY 1015 ---WNSLDEGRHLTNLLSKEIELRNGELQSDYEDATD-----TKPDRDIEEL--S 1062
Db 2406 YLEQLHLQSPKSPFNFKVTIKLPT---LDHTLNETHSYSHKQOSEIDVQTKKQ 2461
QY 1063 ALDTDEPDGQSEPIEELDIOLGISSONDQLNGMAVENGHVPQOQKPEPPKOKK----- 1117
Db 2462 YLTKTKTEASTE-----CSHKQSL-----AERHYQLPKKEKRVTVQL 2498
QY 1118 --QSLGEDHVLBEOKTILPVTSFCSPQPLVPSISNASCLPITTSVSNAGNLIL---KTHVM 1172
Db 2499 PTESIQKNQ---EDKLMQVP-----RKOREFSGSDRGKLPGBSEKQGPSPMIGRKEERLI 2550
QY 1173 SEDKNDFLKPVANGKVV 1189
Db 2551 TERKHEHLKNSAPKVV 2567

; Sequence 126940, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126940
; LENGTH: 1248
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1248)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29438C.1.pap
US-10-437-963-126940

Query Match 3.2%; Score 196.5; DB 16; Length 1248;
Best Local Similarity 21.6%; Pred. No. 0.00044;
Matches 148; Conservative 84; Mismatches 235; Indels 219; Gaps 35;
QY 436 SOEELEKVLRLNKKINATVRTPFLNKGVNNTVTTAGNVISVIGSTETTKGIVPSTNG 495
Db 408 SOGNDXLLMLPRSS-----ILKK---NLTLVLSLSPVQQLIDSAVQIAQSA 458
QY 496 ISNAENSVLSRSTDSLTRALETKYKGVKGVANGQNAAGPSADSV---TENKIGSPPK 552
Db 459 SKSAENSGSQTVPNQEPL---TV-----NNSASTASIALCTPPPLPPPP 503
QY 553 T-----PVSNAATSAGSNVGTLSNVPSQSPFLTRVVPVPHSENIQFQDPRTQIPF 608
Db 504 TVSLAPVSPILPINTSTSIISVLSLSPSPSTAAAPL-PP-----PL 547
QY 609 EVPOYPQTCGYPPPTPVAGVAPCVPRFVRNNVPE-SSLPPASMPYADHYSTFSPDRM 667
Db 548 PPPLKPTVMFPFPPPPPPPPPPPPPP--VSSFSPQPPPPPPSAVP-----GL 593
QY 668 NSSPYQPPPPQYGVPPVPVPSGMVAPVYDSRRIRWPPMYQRDDIIRSNLPPMDVMHSSV 727
Db 594 QASPVPPPPPP--PPMIP-GMKTPP-----TPP-----PPPP----- 624
QY 728 YQTSLRERNYNSLDGYYSVACQPPSEP-----RTTVLPREP-----CGHLKTSCEEQI 775
Db 625 -----AAGQOAPAVPPPPPPPPPPPPPPVPMQTRIPPPPPPSQTNLSVSPFSTS 674
QY 776 RRPQDQWQYHTQKAPLVSSITL-----VATQSPTPSPSPVDFRADFS----- 822
Db 675 KRIPPP-PPPPQTSLSVSS-LPSSRKGNDAAPRPPPPPPPLYSRSHVTSAPSAAPP 732
QY 823 ---VSGTKFEEDHLSHYSP-----WSCGTIGSCINAIIDSEP-----KD 857
Db 733 LPPPKLVGASKPSQEQMITWPPPPPPPKNSNLSLPSKGNVSVSSPPPPPTFSFGAKD 792
QY 858 -VIANSNAVIMDLSDGVKRRVHLFTQRTKKEEDPIIPFSDGPIISKWGAISRSTRGY 916
Db 793 RSTARSRS-----PRSLRPNQSKRT-----PLKPL-----HMKVSR----- 825
QY 917 HTTDPVQATASOGSATKPISVSDVVPVYVNAVQ-SRWSSYGNBATSAAHYVERDFIVTDL 975
Db 826 -----ATQGLWAEQKQKDEASRTPEIDISELESLSFVAMPN---MEEXR----- 867


```
Query Match 3.1%; Score 195; DB 15; Length 2135;
Best Local Similarity 19.1%; Pred. NO. 0.0013;
Matches 254; Conservative 169; Mismatches 528; Indels 378; Gaps 56;

QY 114 GKGVASLNQASALSRMQRKLVTLVNCQLVEBGRVRAARSLGERTVTELLIHOQNP 173
DB 720 GQFSSSLTGVSSQPIHQPOQGIQOQTAPPOQTQVQISQTSSTSEATTQAFVSPQAP 779
QY 174 QQLSANLWAAVRARGCOFLGPA-----MOEALKVLLALEDCSALSRRKVLVLFVORLE 228
DB 780 QVLQVQASQKSTQGVQVAPAEVPAQATQPTTLASSVDSAHSD---VASGMSDGN 836
QY 229 PRFPQAKTSIGHVQVLLYRASCVKYTKRDESSLMOLK-----BEFRSYEARREHDA 282
DB 837 ENVPSSSGRHEGRTTKRHYRKS-VRSRSRHEKTSRPLKAILNVNKNKGDRAVVECOLETHNR 895
QY 283 QIVHIAMEAGLRISQWNSLLY-GDLAKSHMQSIIDKLQSPESPAKSVQELITVLQRT 341
DB 896 KMTVFKFDLD-GDNPEIATIMVNDPILAIERESFVDQVREIIEKADEMLSEDSVSEPE 954
QY 342 GPANLNLRLPHLELLANTID-----PNPDVAVSPTEOLENAMVAVKTVV 385
DB 955 GQGG-----LESQGGKDDYGFSGSKLEGEFKQIIPASSMP--QQIGIPTSSLTQVV 1004
QY 386 HGLVDFIONYSRKGHETPOQPN-----SKYKTSMDCLRLRQGGCPRGTNCTPAHSQ 437
DB 1005 H-----SAGRRFVSPVESRLRESKVPFSEITDTVAASTAQSPG-----MNLSHA 1051
QY 438 ELEKYR-----LRNKKI-----NATVRTFP-----LNKVGNNVTVTTRA----- 473
DB 1052 SLSLQQAQSELRRQMTGEGNTAPNPFSGHTGTFPPVPPFLSSIAAGVPTAAATAPVA 1111
QY 474 -----GNVLSVSGSTTTGKIVPSTNGISNAENS-----VSQILSR 509
DB 1112 TSSPPNDISTSVIQSEVT--VPTREGIAGVATSGVTSGLPIPPVSESPVLSWSWS 1168
QY 510 STDSTLRALETVKVKGVKVGAN-----GQNAAGPSAD-SVTENKIG-----SPKPKPVSNV 558
DB 1169 IIPAVVSISTSPSLOVSTSEIVVSSALYPSVTVAATASAGGSTATGPRPPAVV 1228
QY 559 AATSAGPSNVGTLSNV-----PQKSSPFLTRVVPVPHSENIQYFQ 600
DB 1229 SQQAAGSTTVGATLTSVSTTTPPSTASQLSLSSTSTPTLAETVVVSAHSLD-KTSH 1287
QY 601 DRTQTPFVQVQYQYVPPPP-----TVPAGVPCV-PRFRSNVNPVDESSLPPASMPY 654
DB 1288 SSTTGLAFSL-SAPSSSSPGAGVSSYISQPGGLHPLVPSVIASTPILPQAAGTSTPL 1346
QY 655 ADHYSTFSRDRMNSSPYOPPPQYGPVPPVPSGM-----VAPVYDSR 698
DB 1347 -----LPQVPSIPPLVQPVANVPAVQOOLIHQSQPQALLPNQPHTHCEVDS 1394
QY 699 RIWRPPMYQRDDI-----IRS--NSLPPMDVNHSSV-YQTSILRERYNSLDGYISVA--- 746
DB 1395 T--QPKAPGDDIKTLEKLSLFSEHSSGAQHASVLSLETSTVPGITPAVAP 1452
QY 747 -----CQPPSE-PRTTVPLPREPC---CHLKTSCEBQIRR-KPDQWAOYHTQKAP 791
DB 1453 SKLLTSTTCLPPTNLPLGTVALPVTPVTPVQVSTPVSTTTSVGKPGT---APSKEP 1508
QY 792 LVSS--TLPVATQSPTPPSPLFSDFRADFSVSGTKFEEDHLSHVSPWSCCTGSCINA 850
DB 1509 LTKAPVLPVGTLP-----AGTLFS--EQ 1530
QY 851 IDSEPKDVIANGNAVLMDLDS-----GDVKRRVHLFFETQRRTK----- 888
DB 1531 LPPFPGPSLTQSOQLEDLDAQLRRLTSPEMITVTSAGVPSMAAPTATEACTQPKGV 1590
QY 889 ----EEDPIIFSDGPIISKWG-----AISRSTGYHTDPVQATASQGSATKPISVSD 939
DB 1591 SQVKEGVLTSSGAGVKRGRFQVSAADGAQKKGKNSDEAKSVHFBSSSTSESSVLSS 1650
```

```
QY 940 YVPYNAV-----DSRWSSYGNBATS-SAHYVERDRFIVTDL-----GH 978
DB 1651 SSPESTLVKPEPNGITIPICSSDVPESAHTKTASEAKSDTQPTKVGRFQVTTANKVGR 1710
QY 979 RGHSSGTGLLS-----LELQAKSNSLLQREANALAMQOKWN----- 1016
DB 1711 FSVSKTEKDITDKKEGVPASPPFMDLEQAVLPVPIPKKEKPELSEPSHLNGPSSDPEAA 1770
QY 1017 ----SLDEGR---HUTMLLSKEIELRN-----GELOSDY-TEDATDTKPDRIEELS 1062
DB 1771 FLSDRDVDDGSGPHSPHQLSSKSLPSQLSLSNSFSNYSNDSNEDIEDLKLRL 1830
QY 1063 ALDTEPDQSPSEIIEILDIGLIGSSONDQLN-----GMAVENGHVPQHQKEPPKQ 1115
DB 1831 RL-----RDKHLKEIQDLQSRQKHEIESLYTKLGKVPVAVIIPAAPLSGRRRPTKS 1883
QY 1116 K-----KQSLGEDH-----VILEEQKTLPTVTSFQSPPLVSTSNASCLPI 1156
DB 1884 KGKSSRSRSSLGKNSPQLSGNLGSGAASVLHPQOQLHPPGN-----IPESQNLQLQL 1938
QY 1157 TTSVSAGNL 1165
DB 1939 KPSPSSDNL 1947

RESULT 13
US-10-362-892-9
; Sequence 9, Application US/10362892
; Publication NO. US20040038881A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olga
; APPLICANT: NGUYEN, Daniel B.; WALIA, Narinder K.
; APPLICANT: HAPALIA, April J.A.; YAO, Monique G.
; APPLICANT: GANDHI, Ameena R.; GURURAJAN, Rajagopal
; APPLICANT: DING, Li; PATTERSON, Chandra S.
; APPLICANT: YUE, Henry; BAUGHN, Maria R.
; APPLICANT: TRIBOULEY, Catherine M.; THORNTON, Michael B.
; APPLICANT: ELLIOTT, Vicki S.; LU, Yan
; APPLICANT: ISON, Craig H.; AU-YOUNG, Janice K.
; APPLICANT: TANG, Y. Tom; AZIMZAI, Yalda
; APPLICANT: BURLIGER, John D.; MARCUS, Gregory A.
; APPLICANT: ZINGILL, Kurt A.; LU, Dyung Aina M.
; APPLICANT: LAL, Preeti G.; RAMKUMAR, Jayalaxmi
; APPLICANT: WARREN, Bridget A.; KEARNEY, Liam
; APPLICANT: POLICKY, Jennifer L.; THANGAVELU, Kavitha
; APPLICANT: BURFORD, Neil
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PF-0209 USN
; CURRENT APPLICATION NUMBER: US/10/362,892
; PRIOR FILING DATE: 2003-02-25
; PRIOR FILING DATE: PCT/US01/27219
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,873
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/231,357
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,654
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: US 60/234,902
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US 60/236,499
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: US 60/238,389
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 60/240,542
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PERL Program
; SEQ ID NO 9
; LENGTH: 2135
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040038881A1 5502218CD1
US-10-362-892-9

Query Match      3.1%; Score 195; DB 15; Length 2135;
Best Local Similarity 19.1%; Pred. No. 0.0013;
Matches 254; Conservative 169; Mismatches 528; Indels 378; Gaps 56;

QY 114 GKGVASLNQASLSRPMQKRLVTLNQLVEEGRVRAARSLGERTVTELLIHOQNP 173
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 720 GQFSSSLTGVSSEPIQHPQOQGIQQTAPQOQVQSLQTSQTSSEATTAPQVSOQAP 779
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 174 QQLSANLWAAVRAARGCOFLGPA-----MOEALKVLLALEDCSALSRRKVLVFFVORLE 228
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 780 QVLQVQVAGKSTQGVQVAPAEVPAVQATQPTTLASSVDSAHSD---VASGMSDN 836
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 229 PRFPQASKTSGHVQVLLYRASCVKYTRDEDSLMQLK-----EEFRSYEARREHDA 282
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 837 ENVPSSSGRHEGRTTKRHYRKS-VRSRSRHEKTSRPKRLAINVSNKGDRAVVECOLETHNR 895
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 283 QIVHIAMEAGLRISPRWSSLLY-GDLAKSHMQSIIDKLQSPESPAKSVQBELTVLQRT 341
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 896 KMTVFKFDLD-GDNPEEAIIMVNDPILAIERESFVDQVREIIEKADEMLSEDSVEPE 954
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 342 GQPANLRLRPHLELANID-----PNPDVAVSPTWEQLENAMVAVKTV 385
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 955 GQGG-----LESLOGKDDYGFSGSKLEGEFKQPIPASSMP--QQIGIPTSSLTQVV 1004
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 386 HGLVDFIONYSRKGHETPOQPN-----SKYKTSMCRDLRQOQCGPRGTNCTPAHQ 437
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1005 H-----SAGRRFVSPVESRLRESKVPFSEITDTVAASTAQSPG-----MNLSHA 1051
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 438 ELEKYR-----LANKKI-----NATVTFP-----LLNKVGVNNTVTTA----- 473
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1052 SLSLQQAQFSELRAQMTGPNATPAPNFSGTGFPPVPPFLSSIAGVPTTAATAPVA 1111
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 474 -----GNVLSVIGSTETTKIVPSTNGISNAENS-----VSOLISR 509
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1112 TSPPNDISTSVIQSEVT-----VPTEEGIAGVATSTGVVTSGGLPIDPVSESPVLSVWS 1168
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 510 STDSTLRLALETVKVGVKAN-----GQNAAGPSAD-SVTENKIG-----SPKPTPVSNV 558
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1169 IITPAVVISITTSPLSLQVPTSTSEIVVSSTALYFVTVTSATSASAGSGTATGPKPPAVV 1228
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 559 AATSAGPSNVGTELNSV-----POKSSPFLTRVPVVPVPPHSENIOYFQ 600
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1229 SOQAAGSTTVGATLTSVSTTSPPTASQLSIQLSSSTSTPTLAETVWVVAHSLD-KTSH 1287
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 601 DRPTQIPEEVQYQGYPPRP-----TVPAQVAPCV-PRFVRSNNVPESLPPASMPY 654
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1288 SSTTGAFSL-SAPSSSSGAGVSVYISQPGGLHPLVPSVIASTPILPQAGTSTPL 1346
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 655 ADHYSTFSPDRMNSSPYOPPPQYCPVPPVPSGM-----YAPYDGR 698
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1347 -----LPQVPSIPPLQVPAVPAVQQTLIHSQOPALLPNQPHTHCPREVDSD 1394
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 699 RIWRPMYQRDDI-----IRS--NSLPPMDVMHSSV-YQTSRLRERYNSLDGYVVA--- 746
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1395 T--QKAPGIDDIKTLEELKRLSLFSEHSSGGAHASVLSLTVIESTVTPGPTTAVAP 1452
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 747 -----CQPSB-PRTTVPLPREPC---GHLKATSCBQIIR-KPDQWQHYHTQAP 791
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1453 SKLLTSTTCLPPTNLGLTVALPVTVPVTPQGVSTPVSTTTSGVKPGT----APSKPP 1508
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 792 LVSS--TLPVATQPTPPSPFLSVDFRADFSVSGTKFEBEDHLSHYSPWSCGTIGSCINA 850
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1509 LTKAPVLPVGTCLP-----AGTLPS--BQ 1530
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 851 IDSEPKDVANSNAVLMDLDS-----GDVKRVLHFTQRTK----- 888
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1531 LPPFPGLSTQSOQPLEDLDAQLRLTSLPEMITVTSVAGVPSMAAPTATEAGTPQKGV 1590
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 889 ---EEDPIIFSDGPIISKWG-----AISRSRGTGHTTDPVQATASQGSATKPISVSD 939
```

```
DB 1591 SQVKEGVLATSSGAGVFKMGRFQVSVAAADGAQKGNKSEDAKSVHFSESTSSSVLSS 1650
QY 940 VYPYNAV-----DSRWSSVGNFATS-SAHYVERDRFIVTDL-----CH 978
DB 1651 SSFSESTLVKPEPENGITIFGISDVPESAHTKTASEAKSDTQPTKVGRFQVTTANKVGR 1710
QY 979 RKHSSTGDLSS-----LELOQAKSNSLLLOREANALAMQOKWN----- 1016
DB 1711 FSVSTKTEKIDTKKEGVPASPPFMDLEQAVLPAVIPKKEPELSEPHLNGSPSDPEAA 1770
QY 1017 -----SLDEGR---HLTLNLSKEILRN-----GELQSDY--TEDATDKPRDRIDELELS 1062
DB 1771 FLSDRVDDGSGSPHSPHQLSSKSLPSQNSLSNFSNYSNYSNYSNYSNYSNYSNYSNYSNYS 1830
QY 1063 ALDTEPDQSQSEPIEILDIQLGSSQNDOLLN-----GNVANGHPVQHQKPEPPKQ 1115
DB 1831 RL-----RDKHLKEIQDLOSROKHEIESLYTKLGKVPVPAVIIIPPAAPLSGRRRRPKS 1883
QY 1116 K-----KQSLGEDH-----VILEEQKTLTPVTSQFQPLPVVISNASCLPI 1156
DB 1884 KQSKSRSSSLGKSPQLSGNLSGQSAASVLPQOTLHPGN-----IPESQNLLOLQPL 1938
QY 1157 TTSVSAGNL 1165
DB 1939 KPSPSSDNL 1947

RESULT 14
US-10-196-935A-2
; Sequence 2, Application US/10196935A
; Publication No. US20030082720A1
; GENERAL INFORMATION:
; APPLICANT: Lifton, Richard P
; APPLICANT: Wilson, Frederick H
; APPLICANT: Choate, Keith
; APPLICANT: Ishikawa, Kazuhiko
; APPLICANT: Nelson-Williams, Carole
; TITLE OF INVENTION: COMPOSITIONS METHODS AND KITS RELATING TO TREATING AND DIAGNOSING
; FILE REFERENCE: 044574-5113
; CURRENT APPLICATION NUMBER: US/10/196,935A
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: US 60/306,084
; PRIOR FILING DATE: 2001-07-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 2382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-196-935A-2

Query Match      3.1%; Score 195; DB 14; Length 2382;
Best Local Similarity 18.5%; Pred. No. 0.0015;
Matches 235; Conservative 167; Mismatches 506; Indels 362; Gaps 49;

QY 48 KACPFDTQTAINT-----DIDLVPVNFALLQLVGAQVDPHQSIKLSNLGEN 92
DB 375 KRASPAKSVIGTFEFMAPEMEYEEKYDESVDVYAFGCMLEMATSEYP----- 421
QY 93 KHYEVAKKVEDLALYLKPLSGKGVSALNQSAL-----SRPMOKLVTLVN 139
DB 422 -----YSECONAAQIYRVTVSGK-PASDFKVAIPEVKEIIEGCTIRQNDKERYISKDLN 475
QY 140 COLVBEEGRVRAARSLGERTVTELLIHOQNPQOLSANLWAAVRAARGCOFLGPAQBE 199
DB 476 HAPFQETGVRVLAEEEDDGEKIAIKWLRIEDIKKLK-----GKYKDN 520
QY 200 ALKVLVLAEDGALSRRKVLVFPVQORLEPRFPQASKTSIGHVQVLLYPASCFKTKRDE 259
DB 521 AIEF-----SFDLERDVPEDVAQEMVESGYVCEGDHKTMAKAIKDR--- 561
```


Qy	200	ALKLVLLALEDGSA	SRKVLVLFVVQRLPRFPQ	ASKTSIGHVVOLLVRASC	FKVTKRDE	259
Db	521	ALIEF-----	-----SFDIERDVPED	VAQEWESGYCEGDHKTMA	KADR--	561
Qy	260	DSSLMLQKEFRSYE	ALRREHDAQIVHIAEAG	LRI	SPQWSSLIYGLIAHSHMQSIID	319
Db	562	-VSLIKRKEQR-	-QLVREEQEK-----	-KQESSLKQVQESSASQ	TGIK	604
Qy	320	KLQSPESFAKSVQEL	TIVLQRTGDP-----	ANLRLRPHLELLANID	PNPDVSPTW	371
Db	605	QLPSASTGTPASTT	ASVSTQVEPEPEADQH	QQLQYQPSISVLS--	DGTVD	662
Qy	372	EQLENAMVAVKTVVH	GLVDFIQNYSRKGHET-----	POPOPNKYKTS	MCARDL	419
Db	663	VFTESRVSSQQT	VSIG-----SQEHQAHST	GTVPCHIPSTVQAOSQ	PHGVP	716
Qy	420	ROQGGCPRGCTNCT	FAHSQBELEKYLRLNKK	INATVTPPLLNKVG	VNNVT	479
Db	717	QSQGQ-PSSSSLTG	VSSSQPIQHPQ-----	QOQGIQQTAPQ-----		752
Qy	480	IGSTETTKIVPSTNG	ISNAENSQSOLISRTD	STLRALETVKVKVGK	GANGNAAGPSA	539
Db	753	-----QTVQYLSQ-	TSTSSEAT-----	TAQPV	SQQA	779
Qy	540	DSVTENKIGSPKTP	VSUNVAATSAGPSN--	VGTELNSVPKSSP	FLTRV-----	591
Db	780	PQVLP-QVSAGKOL	PVSPVPTIQGEFQIP	VPATQPSVPVHGAH	FLPVGQPLP	837
Qy	592	HSENIQYFQDPR	TQIPFEVPOX--	POTGYPPPTVPAGV	PCVPRFVR	648
Db	838	-----QYVVSQIP	ISTPHYSTAGTFSS	LPIITWAAGITQ--	PLLTILASSAT	887
Qy	649	-----PASMYPAD	HYSTFSPRDMSSPY-----	QPPP-----	PQYGPVPPVP	694
Db	888	VSTVVP	SQLP-----TLQPVTLQ	PSQVHPQLLQPAV	QSMGIPANLQAAE	942
Qy	695	YDSRRLWRPMPYOR	DDIIRSNSLPPMDVMH	SSVYQTSLEERYNS	LDGYYSVACOP	754
Db	943	YOGFPRLFPQVPG	SNIAPSSNVASVC	IHTSVLSPMPTE	VLATPGYPTTV	1002
Qy	755	TTVPL-----	PREPCGHL-----	KTSCEEQIRKRP	DQMAQVHTQKAP	802
Db	1003	LLVPMGCVGQVQV	SGSLAQAPTSSQ	QAVLES-----	TQVSVQVAPAE	1054
Qy	803	SP--TPPSPL-FS	VDF-----RADFSES	VGTKFEEDHL	SHSPWSCG	851
Db	1055	QPOATQPTTLAS	SVDSAHSDVASGMS	DGNENVPSSSGR	HEGRTTKRHYR	1114
Qy	852	DSEPKDVIAN--	-----SNVLMDL	SGDKVRVHLE	TQRTKEEDPI	906
Db	1115	TSRPKLRILNVN	KNKGRVVECOLETH	NRKWVTFKFDL	DGNPEE	1170
Qy	907	AISRSS-----	RTGYHTTDPVQA-----	TASQ---GSAT	KP	934
Db	1171	ATERESFVQVREI	IEKADEMLSEDS	VEFEGDQGL	ESLQKDDYCF	1230
Qy	935	ISVSDYVPYVNA	VDSRWSSYGN	EATSSAHYVERDR	FI	990
Db	1231	IPASSMFPQIGI	PTS-----SLTQVH	SAGR-RFIVSP	RESRLRESKVP	1282
Qy	991	ELQQA	KSNSLLLOREANALAM	QOKWNSLDEGRHL	TLNLLSKE	1050
Db	1283	AASTAQSGNLSH	SASSLSLOQAFS-----	ELRRAQM----	TEGPNT	1321
Qy	1051	TKPDRD-----	IELELSALD	TDEPDQSEPI	EELD--	1092
Db	1322	APNFSHTGTPF	VPVPFLLSSAGV	PTTAAATAPVAT	SPNDISTSV	1381
Qy	1093	LLNGMAVENGH	FPVQHQKPPQKQSL	GEDHVL	EEQKTI	1152
Db	1382	GIAGVATSTGV	VTSGGLPIPP-----	VSESPVLS	SVSVSITIPAV	1428

Search completed: August 4, 2005, 01:53:32
Job time : 181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:25:48 ; Search time 179 Seconds
(without alignments)
2573.361 Million cell updates/sec

Title: US-10-619-992-2

Perfect score: 6215

Sequence: 1 MPVQAQWTEFLSCPICYNE.....MSEDKNDFLKPANGKMWNS 1191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6215	100.0	1191	5	Aae18608 Human cel
2	6207	99.9	1191	5	Aae18609 Human cel
3	6206	99.9	1191	5	Aae18613 Human cel
4	6206	99.9	1191	5	Aae18612 Human cel
5	6205	99.8	1191	5	Aae18611 Human cel
6	6205	99.8	1191	5	Aae18610 Human cel
7	5483	88.2	1048	7	Adb80934 RING-SH c
8	4416	71.1	869	4	Aam40354 Human pol
9	4066.5	65.4	812	4	Aam40355 Human pol
10	2665.5	42.9	1125	7	Adc37580 Human nuc
11	2577.5	41.5	1109	7	Adf55448 Human nov
12	2106	33.9	421	4	Aag73504 Human gen
13	1593	25.6	306	4	Aam42140 Human gen
14	1593	25.6	306	4	Aam42141 Human pol
15	1538	24.7	819	4	Abb66041 Drosophill
16	1492.5	24.0	296	4	Aag73505 Human gen
17	611	9.8	522	7	Adb65511 Human pro
18	490	7.9	95	5	ABP64303 Human ORF
19	311	5.0	67	5	ABG46802 Human pep
20	267	4.3	276	4	ABG11134 Novel hum
21	206	3.3	1556	8	Adq97964 Human can
22	204.5	3.3	127	3	AAB42232 Human ORF
23	203	3.3	1267	5	ABU65104 Human NOV
24	199.5	3.2	1994	8	Adri8912 Human muc
25	198.5	3.2	3586	7	Adk40970 Novel hum

26	198.5	3.2	3586	8	ADR15689	Adri15689 Kinase 72
27	197.5	3.2	867	8	ADS88438	Ad88438 Human pro
28	197	3.2	713	8	ADR97079	Adr97079 C. albica
29	195.5	3.1	123	4	AAM21438	Aam21438 Peptide #
30	195.5	3.1	123	4	ABB43777	Abb43777 Peptide #
31	195.5	3.1	123	4	AAM37681	Aam37681 Peptide #
32	195.5	3.1	123	4	ABB26719	Abb26719 Protein #
33	195.5	3.1	123	4	AAM77509	Aam77509 Human bon
34	195.5	3.1	123	4	AAM64739	Aam64739 Human bra
35	195.5	3.1	123	4	ABG59147	Abg59147 Human liv
36	195.5	3.1	123	5	ABG46532	Abg46532 Human pep
37	195	3.1	1650	8	ABM81216	Abm81216 Tumour-as
38	195	3.1	2135	5	AAE21714	Aae21714 Human PKI
39	195	3.1	2382	6	ABP71619	Abp71619 Human WNK
40	195	3.1	2382	7	ADJ69598	Adj69598 Human hea
41	195	3.1	2388	8	ADN11119	Adn11119 Human the
42	194	3.1	1267	8	ADN61859	Adn61859 Human nov
43	194	3.1	2185	7	ABO44398	AbO44398 Novel hum
44	194	3.1	2245	7	ABO44382	AbO44382 Novel hum
45	194	3.1	2322	7	ABO44397	AbO44397 Novel hum

ALIGNMENTS

RESULT 1

AAE18608

ID AAE18608 standard; protein; 1191 AA.

AC AAE18608;

DT 17-MAY-2002 (first entry)

DE Human cell surface DNA receptor (DNA-R) #1.

XX Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;
XX otitis media; septic arthritis; drug screening; gene therapy; virucide;
XX antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;
XX auditory.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Domain 14..50

FT Misc-difference /note= "C3HC3D ring finger domain"

FT Misc-difference /note= "N-linked glycosylation"

FT Domain /note= "N-linked glycosylation"

FT Misc-difference /note= "Zinc finger domain"

FT Misc-difference /note= "N-linked glycosylation"

FT Misc-difference /note= "N-linked glycosylation"

FT Misc-difference /note= "N-linked glycosylation"

FT Misc-difference /note= "N-linked glycosylation"

FT Region /note= "N-linked glycosylation"

FT Domain /note= "Proline rich region"

FT Misc-difference /note= "Transmembrane domain"

FT Misc-difference /note= "N-linked glycosylation"

XX WO200210392-A2.

XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024351.

XX 01-AUG-2000; 2000US-0222624P.

PR

XX (UYOR-) UNIV OREGON HEALTH SCI.
PA (USGO) US DEPT VETERANS AFFAIRS.
PA (TARG-) TARGETED GENE DELIVERY.
XX
XX Hefeneider S, Merkins LS, Bennett RM, Seiss DC;
XX WPI: 2002-188747/24.
DR N-PSDB; AAD29592.
XX
PT New mammalian cell surface DNA receptor proteins and nucleic acids,
PT useful for treating inflammation-associated diseases or conditions, e.g.
PT otitis media, septic arthritis, or bacterial or viral infection that
PT causes inflammation.
XX
PS Claim 2; Page 77-82; 94pp; English.
XX
XX The present invention relates to a nucleic acid comprising a sequence
CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble
CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian
CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that
CC effect DNA binding to cells in vivo and in vitro, and for treating
CC inflammation-associated diseases and conditions including otitis media,
CC septic arthritis, and any bacterial or viral infection that causes
CC inflammation by interaction with the DNA-R. These may also be used to
CC screen compounds that modulate binding, uptake and expression. The
CC nucleic acid probes are useful for isolating mammalian species analogues,
CC and for detecting mammalian DNA-R gene expression in cells and tissues.
CC Recombinant expression constructs are useful in molecular biology to
CC transform cells which do not ordinarily express a DNA-R, and the cells
CC are useful as intermediates for making cell membrane preparations for
CC receptor binding assays, which are subsequently useful in drug screening.
CC The recombinant constructs are also useful in gene therapy. The present
CC sequence is human DNA-R. DNA_R gene is located on chromosome 9q34
XX
SQ Sequence 1191 AA;

Query Match 100.0%; Score 6215; DB 5; Length 1191;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVQAAQWTEFSLSCIPYNEFDENVHNPISLSCSHTVCKTCLNKLHRAKCPDQTAINTD 60
DB 1 MPVQAAQWTEFSLSCIPYNEFDENVHNPISLSCSHTVCKTCLNKLHRAKCPDQTAINTD 60
QY 61 IDVLVFNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYLKPLSGGKGVAS 120
DB 61 IDVLVFNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYLKPLSGGKGVAS 120
QY 121 LNQSALSRPMQKLVTLVNCQVLEEGRVRAARAARSLGERTVTELILOHQNPOOLSANL 180
DB 121 LNQSALSRPMQKLVTLVNCQVLEEGRVRAARAARSLGERTVTELILOHQNPOOLSANL 180
QY 181 WAAVARGCQFLGPAMQAEALKLVLALDGSALSRKVLVLFVQVORLEPRFPQAKSTSG 240
DB 181 WAAVARGCQFLGPAMQAEALKLVLALDGSALSRKVLVLFVQVORLEPRFPQAKSTSG 240
QY 241 HVVQLLYRASCFTVKTRDESSLMQKEEFRSYEARLREHDAQIVHIAEAGLRISPEOW 300
DB 241 HVVQLLYRASCFTVKTRDESSLMQKEEFRSYEARLREHDAQIVHIAEAGLRISPEOW 300
QY 301 SSLLYCDLAHKSHMOSIIDKLOSPESFAKSVQELTIVLQRTGDPANLRLPHLELLANI 360
DB 301 SSLLYCDLAHKSHMOSIIDKLOSPESFAKSVQELTIVLQRTGDPANLRLPHLELLANI 360
QY 361 DPNDAVSPTEQLENAMAVKTVVHGLVDFQNTYSRKGHETPQOPNSKYKTSKCRDLR 420
DB 361 DPNDAVSPTEQLENAMAVKTVVHGLVDFQNTYSRKGHETPQOPNSKYKTSKCRDLR 420
QY 421 QGGGCPRGNTCTFAHSQEBELEKYRLRNKKINATVTRTFPLLNKVGNNVTVTAGNVISVI 480
DB 421 QGGGCPRGNTCTFAHSQEBELEKYRLRNKKINATVTRTFPLLNKVGNNVTVTAGNVISVI 480

QY 481 GSTETTGKIVPSTNGISNAENSVSOLISRSRSTDSTLRALETVKYKGVKANGQNAAGPSAD 540
DB 481 GSTETTGKIVPSTNGISNAENSVSOLISRSRSTDSTLRALETVKYKGVKANGQNAAGPSAD 540
QY 541 SVTENKIGSPKTPVSNVAATAGSPSNVGTETNSVPQKSSPFLTRVPYYPHSENIQVFO 600
DB 541 SVTENKIGSPKTPVSNVAATAGSPSNVGTETNSVPQKSSPFLTRVPYYPHSENIQVFO 600
QY 601 DPRTQIPREVPOYPQTYGYYPPPTVPAGVAPCVPRFVRSNNVPSSLPASMPYADHYST 660
DB 601 DPRTQIPREVPOYPQTYGYYPPPTVPAGVAPCVPRFVRSNNVPSSLPASMPYADHYST 660
QY 661 FSPDRMNSSPYQPPPOPGVPVPSGMYAPVYDSRRIRWRPMYQRDDIIRSNLSLPPM 720
DB 661 FSPDRMNSSPYQPPPOPGVPVPSGMYAPVYDSRRIRWRPMYQRDDIIRSNLSLPPM 720
QY 721 DVMSSVYQTSRLRERYNSLDGYYSVACQPPSEPRTTVPLPREPCGHLTKTSEEQIRRPD 780
DB 721 DVMSSVYQTSRLRERYNSLDGYYSVACQPPSEPRTTVPLPREPCGHLTKTSEEQIRRPD 780
QY 781 QMAQYHTOKAPLVSTLPAVATQSTPTPSPLSVDFRADFSVSCTKFEEDHLSHYSPWS 840
DB 781 QMAQYHTOKAPLVSTLPAVATQSTPTPSPLSVDFRADFSVSCTKFEEDHLSHYSPWS 840
QY 841 CGTIGSCINAIIDSEPKDVIANSNVLMDLSDGVKRRVHLFETQRTTKEEDPIIPFSDGP 900
DB 841 CGTIGSCINAIIDSEPKDVIANSNVLMDLSDGVKRRVHLFETQRTTKEEDPIIPFSDGP 900
QY 901 IISKWGAISRSTGYHTTDPVQATASQGSATKPIVSVDYVPYVNAVDSRWSSYGNFATS 960
DB 901 IISKWGAISRSTGYHTTDPVQATASQGSATKPIVSVDYVPYVNAVDSRWSSYGNFATS 960
QY 961 SAHYVERDRFVITDLSGHRKHSSTGDLISLELOAKSNLSLLQREANALAMQOKWNSLDE 1020
DB 961 SAHYVERDRFVITDLSGHRKHSSTGDLISLELOAKSNLSLLQREANALAMQOKWNSLDE 1020
QY 1021 GRHLTLNLLSKIELRNGELQSDYTEDATDKPRDIELELSALDTPDQSGSEPIEIL 1080
DB 1021 GRHLTLNLLSKIELRNGELQSDYTEDATDKPRDIELELSALDTPDQSGSEPIEIL 1080
QY 1081 DIQGISONDQLNGMAVENGHPVQHQKPPPKOKKOSLGEDHVLBEQKTLIPVTSQF 1140
DB 1081 DIQGISONDQLNGMAVENGHPVQHQKPPPKOKKOSLGEDHVLBEQKTLIPVTSQF 1140
QY 1141 SQPLPVSISNASCLPITTSVSGNLIILKTHVMSDKNDFLKPVANGKMWNS 1191
DB 1141 SQPLPVSISNASCLPITTSVSGNLIILKTHVMSDKNDFLKPVANGKMWNS 1191
RESULT 2
AAE18609
ID AAE18609 standard; protein; 1191 AA.
XX
XX AAE18609;
XX AC
XX 17-MAY-2002 (first entry)
XX
XX Human cell surface DNA receptor (DNA-R) #2.
DE Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;
KW otitis media; septic arthritis; drug screening; gene therapy; virucide;
KW antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;
KW auditory.
XX
XX Homo sapiens.
XX
XX WO200210392-A2.
XX
XX 07-FEB-2002.
XX
XX 01-AUG-2001; 2001WO-US024351.
XX
XX 01-AUG-2000; 2000US-0222624P.
PR

XX (UYOR-) UNIV OREGON HEALTH SCI.
PA (USGO) US DEPT VETERANS AFFAIRS.
PA (TARG-) TARGETED GENE DELIVERY.
XX
PI Hefenelder S, Merkins LS, Bennett RM, Seiss DC;
XX
DR WPI; 2002-188747/24.
DR N-PSDB; AAD29593.
XX
XX New mammalian cell surface DNA receptor proteins and nucleic acids, e.g.
PT useful for treating inflammation-associated diseases or conditions, e.g.
PT otitis media, septic arthritis, or bacterial or viral infection that
PT causes inflammation.
XX
XX Disclosure; Page 88-94; 94pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble
CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian
CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that
CC effect DNA binding to cells in vivo and in vitro, and for treating
CC inflammation-associated diseases and conditions including otitis media,
CC septic arthritis, and any bacterial or viral infection that causes
CC inflammation by interaction with the DNA-R. These may also be used to
CC screen compounds that modulate binding, uptake and expression. The
CC nucleic acid probes are useful for isolating mammalian species analogues,
CC and for detecting mammalian DNA-R gene expression in cells and tissues.
CC Recombinant expression constructs are useful in molecular biology to
CC transform cells which do not ordinarily express a DNA-R, and the cells
CC are useful as intermediates for making cell membrane preparations for
CC receptor binding assays, which are subsequently useful in drug screening.
CC The recombinant constructs are also useful in gene therapy. The present
CC sequence is human DNA-R. DNA_R gene is located on chromosome 9q34
XX
SQ Sequence 1191 AA;

Query Match 99.9%; Score 6207; DB 5; Length 1191;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MPVQAQWTEFLSCPICYNEFDENVHKKPISLGCSTVCKTCLNKLHKKACPFDDTAINTD 60
DB 1 MAVQAQWTEFLSCPICYNEFDENVHKKPISLGCSTVCKTCLNKLHKKACPFDDTAINTD 60

QY 61 IDVLPVNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKCVEDLALYKPLSGGKGVAS 120
DB 61 IDVLPVNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKCVEDLALYKPLSGGKGVAS 120

QY 121 LQNSALSRPQWOKLVTLVNCQVVEEGRVRAMRAARSIGERTVTTELILQHONPQQLSANL 180
DB 121 LQNSALSRPQWOKLVTLVNCQVVEEGRVRAMRAARSIGERTVTTELILQHONPQQLSANL 180

QY 181 WAAVRAGCQFLGPMQAEALKLVLLALEDDGSALSRLKVLVLFVQRLPRFPQASKTSG 240
DB 181 WAAVRAGCQFLGPMQAEALKLVLLALEDDGSALSRLKVLVLFVQRLPRFPQASKTSG 240

QY 241 HVVQLLYRASCFVKTKRDESSLMQLKEEFSRYSYALRRHDAQIVHIAEAGLIRISPEQW 300
DB 241 HVVQLLYRASCFVKTKRDESSLMQLKEEFSRYSYALRRHDAQIVHIAEAGLIRISPEQW 300

QY 301 SLLYGLAHKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRPHLELLANI 360
DB 301 SLLYGLAHKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRPHLELLANI 360

QY 361 DPNPDVSPTEQLENAMAVKTVVHGLVDFIQNSRKGHETPQOPNSKYKTSKMCRLR 420
DB 361 DPNPDVSPTEQLENAMAVKTVVHGLVDFIQNSRKGHETPQOPNSKYKTSKMCRLR 420

QY 421 QQGGCPRGNTCTFAHSEEELEKYLRLNKKINATVTRTFPLLNKVGNNVTVTAGNVISVI 480
DB 421 QQGGCPRGNTCTFAHSEEELEKYLRLNKKINATVTRTFPLLNKVGNNVTVTAGNVISVI 480

QY 481 GSTETTKIVPSTNGISNAENSVSGLISRSTDTSTLRALETVKVKGVKGVANGQNAAGPSAD 540
DB 481 GSTETTKIVPSTNGISNAENSVSGLISRSTDTSTLRALETVKVKGVKGVANGQNAAGPSAD 540

QY 541 SVTENKIGSPKTPVSNVAATSAAGPSNVGTENSVPOKSSPDLTRVPVYPHSENIQYFQ 600
DB 541 SVTENKIGSPKTPVSNVAATSAAGPSNVGTENSVPOKSSPDLTRVPVYPHSENIQYFQ 600

QY 601 DPRTQIPPEVQYQYQYPPPTVPAGVPCVPRFVRSNVNPVPSLPASMPYADHYST 660
DB 601 DPRTQIPPEVQYQYQYPPPTVPAGVPCVPRFVRSNVNPVPSLPASMPYADHYST 660

QY 661 FSPDRMNSSPYQPPPPQYGPVPGMVAIVYDSRRIRWPPMYQRDDIIRSNLSLPM 720
DB 661 FSPDRMNSSPYQPPPPQYGPVPGMVAIVYDSRRIRWPPMYQRDDIIRSNLSLPM 720

QY 721 DVMSVSVQTSIRERYNSLDGYVSACOPPSPRITVPLPREPCGHLTKTSCSEQIRRKPD 780
DB 721 DVMSVSVQTSIRERYNSLDGYVSACOPPSPRITVPLPREPCGHLTKTSCSEQIRRKPD 780

QY 781 QWAOYHTQKAPLVSSSTLPVATQSPPTPSPPLFSVDVFRADFSESVSGTKFEEDHLSHYSPWS 840
DB 781 QWAOYHTQKAPLVSSSTLPVATQSPPTPSPPLFSVDVFRADFSESVSGTKFEEDHLSHYSPWS 840

QY 841 COTIGSCINADSEPKDVIANSNVLMOLDGSDVKRRVHLFETQRTKEEDPIIPPSDGP 900
DB 841 COTIGSCINADSEPKDVIANSNVLMOLDGSDVKRRVHLFETQRTKEEDPIIPPSDGP 900

QY 901 IISKWGAISRSTGYHTTDPVQATASOGSATKPIVSVDYVYVNAVDSRWSYGNATS 960
DB 901 IISKWGAISRSTGYHTTDPVQATASOGSATKPIVSVDYVYVNAVDSRWSYGNATS 960

QY 961 SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQQAQKSNLQLQREANALAMQKNSLDE 1020
DB 961 SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQQAQKSNLQLQREANALAMQKNSLDE 1020

QY 1021 GRHLTLNLSKEIELRNGELQSDYTEDATDTKPRDIELELSALDTPDQSGSEPIEIL 1080
DB 1021 GRHLTLNLSKEIELRNGELQSDYTEDATDTKPRDIELELSALDTPDQSGSEPIEIL 1080

QY 1081 DIQLGISSNDOLLNGMAVENGHVQHQKEPKKQKQSLGSDHVLLEOKTILPVTSCF 1140
DB 1081 DIQLGISSNDOLLNGMAVENGHVQHQKEPKKQKQSLGSDHVLLEOKTILPVTSCF 1140

QY 1141 SLPPLVSIASCLPITTSVSNAGNLLKTHVMSDKNDLKPVANGKMYNS 1191
DB 1141 SLPPLVSIASCLPITTSVSNAGNLLKTHVMSDKNDLKPVANGKMYNS 1191

RESULT 3
AAE18613
ID AAE18613 standard; protein; 1191 AA.
XX
AC AAE18613;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human cell surface DNA receptor (DNA-R) #1 mutant, C431A.
XX
KW Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;
KW otitis media; septic arthritis; drug screening; gene therapy; virucide;
KW antinflammatory; antibacterial; immunosuppressive; chromosome 9q34;
XX
XX auditory; mutant; mutein.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT Misc-difference 431 /note= "wild type Cys substituted with Ala"
XX
XX WO200210392-A2.
XX
XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024351.
 XX 01-AUG-2000; 2000US-0222624P.
 PR (UYOR-) UNIV OREGON HEALTH SCI.
 PA (USGO) US DEPT VETERANS AFFAIRS.
 PA (TARG-) TARGETED GENE DELIVERY.
 XX Hefeneider S, Merkins LS, Bennett RM, Seiss DC;
 XX WPI; 2002-188747/24.
 XX New mammalian cell surface DNA receptor proteins and nucleic acids,
 PT useful for treating inflammation-associated diseases or conditions, e.g.
 PT otitis media, septic arthritis, or bacterial or viral infection that
 PT causes inflammation.
 XX Example 2; Page; 94pp; English.
 XX The present invention relates to a nucleic acid comprising a sequence
 CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble
 CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian
 CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that
 CC effect DNA binding to cells in vivo and in vitro, and for treating
 CC inflammation-associated diseases and conditions including otitis media,
 CC septic arthritis, and any bacterial or viral infection that causes
 CC inflammation by interaction with the DNA-R. These may also be used to
 CC screen compounds that modulate binding, uptake and expression. The
 CC nucleic acid probes are useful for isolating mammalian species analogues,
 CC and for detecting mammalian DNA-R gene expression in cells and tissues.
 CC Recombinant expression constructs are useful in molecular biology to
 CC transform cells which do not ordinarily express a DNA-R, and the cells
 CC are useful as intermediates for making cell membrane preparations for
 CC receptor binding assays, which are subsequently useful in drug screening.
 CC The recombinant constructs are also useful in gene therapy. The present
 CC sequence is human DNA-R mutant. DNA R gene is located on chromosome 9q34.
 CC Note: The present sequence is not shown in the specification, but is
 CC derived from the wild type DNA-R sequence shown in sequence listing (page
 CC 77-82) of the specification
 XX Sequence 1191 AA;
 Query Match 99.9%; Score 6206; DB 5; Length 1191;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MPVQAQWTEFLSCPICYNEFDENVHVKPISLGCSTVCKTCLNKLHRRKACPPDQTAINTD 60
 DB 1 MPVQAQWTEFLSCPICYNEFDENVHVKPISLGCSTVCKTCLNKLHRRKACPPDQTAINTD 60
 QY 61 IDVLVFNFALLQVGAQVPDQHSIKLSNLGENKHVEAVKCVEDIALYLKPLSGKGVAS 120
 DB 61 IDVLVFNFALLQVGAQVPDQHSIKLSNLGENKHVEAVKCVEDIALYLKPLSGKGVAS 120
 QY 121 LNQSALSRPMORLVTNLCQVVEEGRVRAARARSIGERTVTELLIHOQNPOQLSANL 180
 DB 121 LNQSALSRPMORLVTNLCQVVEEGRVRAARARSIGERTVTELLIHOQNPOQLSANL 180
 QY 181 WAAVRAQCQFLGPAQVEALXVLLALEDSALSARKVLVLFVQVORLEPRFPQASKTSIG 240
 DB 181 WAAVRAQCQFLGPAQVEALXVLLALEDSALSARKVLVLFVQVORLEPRFPQASKTSIG 240
 QY 241 HVVQLLYRASCPCVKTKRDESSLMQLKBEFRSYEARREHDAQIVHIAEAGLRISPEQW 300
 DB 241 HVVQLLYRASCPCVKTKRDESSLMQLKBEFRSYEARREHDAQIVHIAEAGLRISPEQW 300
 QY 301 SLLYGDIAHKSHMOSIIDKLOSPSFAKSVQELTIVLQRTGDPANLRLPHELLANI 360
 DB 301 SLLYGDIAHKSHMOSIIDKLOSPSFAKSVQELTIVLQRTGDPANLRLPHELLANI 360
 QY 361 DPNDAVSPTEQLENAMVAVKTVVHGLVDFIQNYSRKGHETPQPQPNISKYKTSRCRDLR 420

DB DPNDAVSPTEQLENAMVAVKTVVHGLVDFIQNYSRKGHETPQPQPNISKYKTSRCRDLR 420
 QY 421 OQGCPRGTCTFAHSEEELEKYRLRNKKINATVTRTPLLNKGVNNTVTTTGNVSVI 480
 DB 421 OQGCPRGTCTFAHSEEELEKYRLRNKKINATVTRTPLLNKGVNNTVTTTGNVSVI 480
 QY 481 GSTETTGKIVPSTNGISNAENSVSQLSRSRSDTLRALETVKYKGVKGANGQNAAGPSAD 540
 DB 481 GSTETTGKIVPSTNGISNAENSVSQLSRSRSDTLRALETVKYKGVKGANGQNAAGPSAD 540
 QY 541 SVTENKIGSPKPTVSNVAATSAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQIFQ 600
 DB 541 SVTENKIGSPKPTVSNVAATSAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSENIQIFQ 600
 QY 601 DPTQIPEVPOYQOTGYPPPTVPAGVPCVPRFVRSNNVPSSLPASPMPADHYST 660
 DB 601 DPTQIPEVPOYQOTGYPPPTVPAGVPCVPRFVRSNNVPSSLPASPMPADHYST 660
 QY 661 FSPDRWNSSPYQPPQYGVPPVPSGMYPVYDSRRIRWRPPMYQRDDIIRSNLPPM 720
 DB 661 FSPDRWNSSPYQPPQYGVPPVPSGMYPVYDSRRIRWRPPMYQRDDIIRSNLPPM 720
 QY 721 DVMSSVYQTSRLRYSNLDGYYSVACOPPSEPRTTVPDPCHLTKTSCCEQIRRPD 780
 DB 721 DVMSSVYQTSRLRYSNLDGYYSVACOPPSEPRTTVPDPCHLTKTSCCEQIRRPD 780
 QY 781 QMAQYHTQKAPLVGSTLPVATQSPTPSPFSDVFRADFSVSCTKFEEDHLSHSPWS 840
 DB 781 QMAQYHTQKAPLVGSTLPVATQSPTPSPFSDVFRADFSVSCTKFEEDHLSHSPWS 840
 QY 841 CGTIGSCINADSPKDVIANSNVLMDLSDGDVKKRVHLFETQRTTKEEDPIIPFSDGP 900
 DB 841 CGTIGSCINADSPKDVIANSNVLMDLSDGDVKKRVHLFETQRTTKEEDPIIPFSDGP 900
 QY 901 IISKWALSRSSRTGYHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSSYGNEATS 960
 DB 901 IISKWALSRSSRTGYHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSSYGNEATS 960
 QY 961 SAHYVERDRFTVDTLSGHRKHSSTGDLISLELQAKSNLLOREANALAMQCKWNSLDE 1020
 DB 961 SAHYVERDRFTVDTLSGHRKHSSTGDLISLELQAKSNLLOREANALAMQCKWNSLDE 1020
 QY 1021 GRHLTLNLLSKEIELRNGELQSDYTEDATDKPORDIELELSALDTPDQSQSEPIEIL 1080
 DB 1021 GRHLTLNLLSKEIELRNGELQSDYTEDATDKPORDIELELSALDTPDQSQSEPIEIL 1080
 QY 1081 DIQLGISQNDQLLNGMAVENGHPVQHQKPPKQKQSLGSDHVLBEQKTLTPVTSQF 1140
 DB 1081 DIQLGISQNDQLLNGMAVENGHPVQHQKPPKQKQSLGSDHVLBEQKTLTPVTSQF 1140
 QY 1141 SQPLPVSTSNASCUPITTSVSAGNLIILKTHVMSDEKNDPLKPVANGKMWNS 1191
 DB 1141 SQPLPVSTSNASCUPITTSVSAGNLIILKTHVMSDEKNDPLKPVANGKMWNS 1191
 RESULT 4
 AAEL18612
 ID AAEL18612 standard; protein; 1191 AA.
 XX
 AC AAEL18612;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human cell surface DNA receptor (DNA-R) #1 mutant, C416A.
 XX
 KW Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;
 KW otitis media; septic arthritis; drug screening; gene therapy; virucide;
 KW antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;
 KW auditory; mutant; muten.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers

FT Misc-difference 416 /note= "wild type Cys substituted with Ala"
 XX PT
 XX
 PN WO200210392-A2.
 XX
 PD 07-FEB-2002.
 XX
 PF 01-AUG-2001; 2001WO-US024351.
 XX
 PR 01-AUG-2000; 2000US-0222624P.
 XX
 PA (UYOR-) UNIV OREGON HEALTH SCI.
 PA (USGO) US DEPT VETERANS AFFAIRS.
 PA (TARG-) TARGETED GENE DELIVERY.
 XX
 PI Hefeneider S, Merkins LS, Bennett RM, Seiss DC;
 XX
 DR WPI; 2002-188747/24.
 XX
 XX New mammalian cell surface DNA receptor proteins and nucleic acids,
 PT useful for treating inflammation-associated diseases or conditions, e.g.
 PT otitis media, septic arthritis, or bacterial or viral infection that
 PT causes inflammation.
 XX
 XX Example 2; Page; 94pp; English.
 PS
 XX
 CC The present invention relates to a nucleic acid comprising a sequence
 CC encoding a mammalian cell surface DNA receptor (DNA-R), a soluble
 CC mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian
 CC DNA-R. Mammalian DNA-R proteins may be used to screen compounds that
 CC effect DNA binding to cells in vivo and in vitro, and for treating
 CC inflammation-associated diseases and conditions including otitis media,
 CC septic arthritis, and any bacterial or viral infection that causes
 CC inflammation by interaction with the DNA-R. These may also be used to
 CC screen compounds that modulate binding, uptake and expression. The
 CC nucleic acid probes are useful for isolating mammalian species analogues,
 CC and for detecting mammalian DNA-R gene expression in cells and tissues.
 CC Recombinant expression constructs are useful in molecular biology to
 CC transform cells which do not ordinarily express a DNA-R, and the cells
 CC are useful as intermediates for making cell membrane preparations for
 CC receptor binding assays, which are subsequently useful in drug screening.
 CC The recombinant constructs are also useful in gene therapy. The present
 CC sequence is human DNA-R mutant. DNA R gene is located on chromosome 9q34.
 CC Note: The present sequence is not shown in the specification, but is
 CC derived from the wild type DNA-R sequence shown in sequence listing (page
 CC 77-82) of the specification
 XX
 SQ Sequence 1191 AA;
 Query Match 99.9%; Score 6206; DB 5; Length 1191;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MPVQAQWTEFLSCPICYNEFDENVHVKPSISLGSCHTVCKTCLNKLHAKCPDQTAINTD 60
 DB 1 MPVQAQWTEFLSCPICYNEFDENVHVKPSISLGSCHTVCKTCLNKLHAKCPDQTAINTD 60
 QY 61 IDVLPNVFNALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVBDDLALYLKPLSGGKGVAS 120
 DB 61 IDVLPNVFNALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVBDDLALYLKPLSGGKGVAS 120
 QY 121 LNQSALSRPMQKLVTLVNCQVVEEGRVRAAARSLGERTVTELIQHONPQQLSANTL 180
 DB 121 LNQSALSRPMQKLVTLVNCQVVEEGRVRAAARSLGERTVTELIQHONPQQLSANTL 180
 QY 181 WAAVARGCQFLGPAQOEALKLVLALDGSALSRRKVLVLFVQRLPRFPQASKTSIG 240
 DB 181 WAAVARGCQFLGPAQOEALKLVLALDGSALSRRKVLVLFVQRLPRFPQASKTSIG 240
 QY 241 HVVQLLYRASCFFVTKRDESSLMQLKEEPRSYEARLRREHDAQIVHIAEAGLRISPEQW 300
 DB 241 HVVQLLYRASCFFVTKRDESSLMQLKEEPRSYEARLRREHDAQIVHIAEAGLRISPEQW 300

QY 301 SSSLYGDLAHKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRPHLELLANI 360
 DB 301 SSSLYGDLAHKSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRPHLELLANI 360
 QY 361 DNPDAVSPPTWEQLENAMVAVKTVVHGLVDFIQTNSRKGHETPOQPNKYKTSWCRDLR 420
 DB 361 DNPDAVSPPTWEQLENAMVAVKTVVHGLVDFIQTNSRKGHETPOQPNKYKTSWCRDLR 420
 QY 421 QQGGCPRGTCTCTFAHSQBELEKYLRLNKKINATVTFPLLNKVGYNNTVTTTGNVSVI 480
 DB 421 QQGGCPRGTCTCTFAHSQBELEKYLRLNKKINATVTFPLLNKVGYNNTVTTTGNVSVI 480
 QY 481 GSTETTGKIVPSTNGISNAENSVSQIISRSTDSITLRALETVKVKGVANGONAAQPSAD 540
 DB 481 GSTETTGKIVPSTNGISNAENSVSQIISRSTDSITLRALETVKVKGVANGONAAQPSAD 540
 QY 541 SVTENKIGSPPKTPVSNVAATSAGPSNVGTSLNSVPQKSSPELTVPVVPVPHSENIOYFQ 600
 DB 541 SVTENKIGSPPKTPVSNVAATSAGPSNVGTSLNSVPQKSSPELTVPVVPVPHSENIOYFQ 600
 QY 601 DPRTQIPPEVPOYPOTGYPPPPPTVPAGVAPCVPRFVRSNNVPESSLPASMPYADHYST 660
 DB 601 DPRTQIPPEVPOYPOTGYPPPPPTVPAGVAPCVPRFVRSNNVPESSLPASMPYADHYST 660
 QY 661 FSPDRMNSSPYQPPPPQPPYGPVPPVPSGMYAPVTDSDRRIWRPMPYQRODDIIRNSLPPM 720
 DB 661 FSPDRMNSSPYQPPPPQPPYGPVPPVPSGMYAPVTDSDRRIWRPMPYQRODDIIRNSLPPM 720
 QY 721 DYMHSVVQTSRLRERYNSLDGYVACQPPSPRTTVPPLPRPCGHLTKSCREQIRRKPD 780
 DB 721 DYMHSVVQTSRLRERYNSLDGYVACQPPSPRTTVPPLPRPCGHLTKSCREQIRRKPD 780
 QY 781 QMAQYHTOKAPLVSSSTLPVATQSPPTPSPLFSDVDFRADPSESVSCTKFBEDHLSHYSPMS 840
 DB 781 QMAQYHTOKAPLVSSSTLPVATQSPPTPSPLFSDVDFRADPSESVSCTKFBEDHLSHYSPMS 840
 QY 841 CGTIGSCINAIDSEPKDVIANSNVLMOLDSDGDKRRVHLFTQRTKESDPIIPFSDGP 900
 DB 841 CGTIGSCINAIDSEPKDVIANSNVLMOLDSDGDKRRVHLFTQRTKESDPIIPFSDGP 900
 QY 901 IISKWGAISRRTGYHTTDPVQATASQGSATKPIVSVDYVYVNVAVDSRWSSYGNATS 960
 DB 901 IISKWGAISRRTGYHTTDPVQATASQGSATKPIVSVDYVYVNVAVDSRWSSYGNATS 960
 QY 961 SAHYVERDRFIYTDLSGHRKHSSTGDLISLELQQAQKSNLLQREANALAMQKNSLDE 1020
 DB 961 SAHYVERDRFIYTDLSGHRKHSSTGDLISLELQQAQKSNLLQREANALAMQKNSLDE 1020
 QY 1021 GRHLTLNLSKEIELRNGELQSDYTDATDTTKPDRDIELELSALDTPDQSQSEPIEIL 1080
 DB 1021 GRHLTLNLSKEIELRNGELQSDYTDATDTTKPDRDIELELSALDTPDQSQSEPIEIL 1080
 QY 1081 DIQLGISSNDOLLNGMAVENGHVPQHQKEPPKQKQSLGHDHVILEQKTLIPVTSF 1140
 DB 1081 DIQLGISSNDOLLNGMAVENGHVPQHQKEPPKQKQSLGHDHVILEQKTLIPVTSF 1140
 QY 1141 SQPLPVSISNASCLPITTSVSAGNLILKTHVMASEDKNDFLKPVANGKWN 1191
 DB 1141 SQPLPVSISNASCLPITTSVSAGNLILKTHVMASEDKNDFLKPVANGKWN 1191
 RESULT 5
 AAE18611
 ID AAE18611 standard; protein; 1191 AA.
 XX AAE18611;
 XX AC
 XX XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human cell surface DNA receptor (DNA-R) #1 mutant, C431S.
 XX Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;
 KW otitis media; septic arthritis; drug screening; gene therapy; virucide;

DT	17-MAY-2002	(first entry)	
XX	Human cell surface DNA receptor (DNA-R) #1 mutant, C416S.		
XX	Human; cell surface DNA receptor; DNA-R; inflammation-associated disease; otitis media; septic arthritis; drug screening; gene therapy; virucide; antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34; auditory; mutant; mutein.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 416	/note= "Wild type Cys substituted with Ser"	
FT	WO200210392-A2.		
XX	07-FEB-2002.		
XX	01-AUG-2001; 2001WO-US024351.		
XX	01-AUG-2000; 2000US-0222624P.		
XX	(UYOR-) UNIV OREGON HEALTH SCI.		
PA	(USGO) US DEPT VETERANS AFFAIRS.		
PA	(TARG-) TARGETED GENE DELIVERY.		
XX	Hefeneider S, Merkins LS, Bennett RM, Seis DC;		
PI	WPI; 2002-198747/24.		
XX	New mammalian cell surface DNA receptor proteins and nucleic acids, useful for treating inflammation-associated diseases or conditions, e.g. otitis media, septic arthritis, or bacterial or viral infection that causes inflammation.		
XX	Example 2; Page: 94pp; English.		
XX	The present invention relates to a nucleic acid comprising a sequence encoding a mammalian cell surface DNA receptor (DNA-R), a soluble mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian DNA-R. Mammalian DNA-R proteins may be used to screen compounds that effect DNA binding to cells in vivo and in vitro, and for treating inflammation-associated diseases and conditions including otitis media, septic arthritis, and any bacterial or viral infection that causes inflammation by interaction with the DNA-R. These may also be used to screen compounds that modulate binding, uptake and expression. The nucleic acid probes are useful for isolating mammalian species analogues, and for detecting mammalian DNA-R gene expression in cells and tissues. Recombinant expression constructs are useful in molecular biology to transform cells which do not ordinarily express a DNA-R, and the cells are useful as intermediates for making cell membrane preparations for receptor binding assays, which are subsequently useful in drug screening. The recombinant constructs are also useful in gene therapy. The present sequence is human DNA-R mutant. DNA R gene is located on chromosome 9q34. Note: The present sequence is not shown in the specification, but is derived from the wild type DNA-R sequence shown in sequence listing (page 77-82) of the specification		
XX	Sequence 1191 AA;		
XX	Query Match 99.8%; .Score 6205; DB 5; Length 1191;		
XX	Best Local Similarity 99.9%; Pred. No. 0;		
XX	Matches 1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1 MPVQAQWTEFLSCPICTYNEFDENVHKPISLGCSTVCKTCLNKLHRKACPFDDQTAINTD 60		
DB	1 MPVQAQWTEFLSCPICTYNEFDENVHKPISLGCSTVCKTCLNKLHRKACPFDDQTAINTD 60		
QY	61 IDVLPNVFNALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVEDLALYLKPLSGGKGVAS 120		
DB	61 IDVLPNVFNALLQLVGAQVDPHQSIKLSNLGENKHVEYAKKCVEDLALYLKPLSGGKGVAS 120		

QY	121	LNQSALSRPMQKLVTLVNCQVBEGRVNRAMRAARSLGERTVTTELLILQHQNPOOLSANL 180
DB	121	LNQSALSRPMQKLVTLVNCQVBEGRVNRAMRAARSLGERTVTTELLILQHQNPOOLSANL 180
QY	181	WAAVARGCOFLGPAMQEEALKVLVALEDGALSARKVLVLFVQORLPRFPQASKTSIG 240
DB	181	WAAVARGCOFLGPAMQEEALKVLVALEDGALSARKVLVLFVQORLPRFPQASKTSIG 240
QY	241	HVVQLLYRASCFCVTKRDEDSLSMLQKEEFRSYEALRREHDAQIVHIAEAGLRISPEQW 300
DB	241	HVVQLLYRASCFCVTKRDEDSLSMLQKEEFRSYEALRREHDAQIVHIAEAGLRISPEQW 300
QY	301	SSLLYGDIAHSHMQSIIDKLQSPESPAKSVOELTIVLQRTGCDPANLRLRPHLELLANI 360
DB	301	SSLLYGDIAHSHMQSIIDKLQSPESPAKSVOELTIVLQRTGCDPANLRLRPHLELLANI 360
QY	361	DNPDVAVSPTWEOLENAMVAVKTVVHGLVDFIQTNSRKGHETPQPPNSKYKTSRCRDLR 420
DB	361	DNPDVAVSPTWEOLENAMVAVKTVVHGLVDFIQTNSRKGHETPQPPNSKYKTSRCRDLR 420
QY	421	QOQGCPRGTNCTFAHSQBELEKYLRLNKKINATVTRFPLLNKVGYNVNTVTTTGNVISVI 480
DB	421	QOQGCPRGTNCTFAHSQBELEKYLRLNKKINATVTRFPLLNKVGYNVNTVTTTGNVISVI 480
QY	481	GSTETTKIVPSTNGISNAENSVSQIISRSTDSLRALETVKVKGKVGANGQNAAGPSAD 540
DB	481	GSTETTKIVPSTNGISNAENSVSQIISRSTDSLRALETVKVKGKVGANGQNAAGPSAD 540
QY	541	SVTENKISGPKTPVSNVAATSAGPSNVGTSLNSVPQKSSPPLTRVPPVPHSENIQYFQ 600
DB	541	SVTENKISGPKTPVSNVAATSAGPSNVGTSLNSVPQKSSPPLTRVPPVPHSENIQYFQ 600
QY	601	DPRTQIPEVQYQYPTGYPPPTVPAGVAPCVPRFVRSNNVPESSLPASMPYADHYST 660
DB	601	DPRTQIPEVQYQYPTGYPPPTVPAGVAPCVPRFVRSNNVPESSLPASMPYADHYST 660
QY	661	FSPRDRMNSSPYQPPPPQYGPVPPVPSGMVAPVTDSSRIWRPMPYQRDDIIIRNSLPPM 720
DB	661	FSPRDRMNSSPYQPPPPQYGPVPPVPSGMVAPVTDSSRIWRPMPYQRDDIIIRNSLPPM 720
QY	721	DMVHSSVQTSIRERYNSLDGYSVACOPPEPRTTVPLPREPCGHLTKSCSEQIRRKPD 780
DB	721	DMVHSSVQTSIRERYNSLDGYSVACOPPEPRTTVPLPREPCGHLTKSCSEQIRRKPD 780
QY	781	QWAAVHTQKAPLVSSSTLPVATQSPPTPPLFSVDPRADFSVSGTKFEDHLSHYSPWS 840
DB	781	QWAAVHTQKAPLVSSSTLPVATQSPPTPPLFSVDPRADFSVSGTKFEDHLSHYSPWS 840
QY	841	CGTIGSCINADSEPKDVIANSNVLMDLDSGDVKRRVHLFETQRTKEEDPIIPPSDGP 900
DB	841	CGTIGSCINADSEPKDVIANSNVLMDLDSGDVKRRVHLFETQRTKEEDPIIPPSDGP 900
QY	901	IISKVGAISRRTGYHTTDPVQATASQGSATKPIVSNDYVPYVNAVDSRWSYNGEATS 960
DB	901	IISKVGAISRRTGYHTTDPVQATASQGSATKPIVSNDYVPYVNAVDSRWSYNGEATS 960
QY	961	SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQQAQKSNLILQREANALAMQKWNLSDE 1020
DB	961	SAHYVERDRFIVTDLSGHRKHSSTGDLISLELQQAQKSNLILQREANALAMQKWNLSDE 1020
QY	1021	GRHLTLNLSKEIELRNGELQSDYTETADTDTKPRDRIEELSALDTDEPDGQSEPIEIL 1080
DB	1021	GRHLTLNLSKEIELRNGELQSDYTETADTDTKPRDRIEELSALDTDEPDGQSEPIEIL 1080
QY	1081	DTQLGISNDOLLNGMAVENCHPVQHQKEPPKQKQSLGSDHVTLEQKTLIPVTSQCF 1140
DB	1081	DTQLGISNDOLLNGMAVENCHPVQHQKEPPKQKQSLGSDHVTLEQKTLIPVTSQCF 1140
QY	1141	SOPLPVVISNASCLPITTSVSAGNLLILKTHVMSDKNDKDFLKPVANGKMNVS 1191
DB	1141	SOPLPVVISNASCLPITTSVSAGNLLILKTHVMSDKNDKDFLKPVANGKMNVS 1191

RESULT 7
ADB80934
ID ADB80934 standard; protein; 1048 AA.
XX
AC ADB80934;
XX
DT 04-DEC-2003 (first entry)
XX
DE RING-SH complex related protein, SEQ ID No 8.
XX
KW RING-SH 3; Gag protein; Gag late domain; PI3K; actin; myosin; Hsp60;
KW Hep70; Hsp90; STAM1; STAM2A; STAM2B; VHS-UIM; GTPase; E2 enzyme; teg101;
KW cullin; RING-SH; clathrin; virucide; vaccine; antiviral; retrovirus;
KW rhabdovirus; filovirus.
XX
OS Unidentified.
XX
PN WO2003033646-A2.
XX
PD 24-APR-2003.
XX
PF 31-JUL-2002; 2002WO-US024589.
XX
PR 31-JUL-2001; 2001US-0308958P.
PR 09-NOV-2001; 2001US-0345846P.
XX
PA (PROT-) PROTEOLOGICS INC.
XX
PI Greener T, Moskowitz H, Reiss Y, Alroy I;
XX
DR WPI: 2003-393509/37.
DR N-PSDB; ADB80975.
XX
PT New isolated protein complex comprising a RING-SH 3 polypeptide and
PT another polypeptide, useful for detecting cells infected with a virus,
PT and for treating viral disorders caused by retroviruses, rhabdoviruses,
PT or filoviruses.
XX
XX Disclosure; Fig 8; 176pp; English.
XX
XX The invention relates to a novel isolated protein complex comprising a
CC RING-SH 3 polypeptide and a polypeptide selected from a Gag protein, a
CC Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A,
CC STAM2B, VHS-UIM, a GTPase, an E2 enzyme, teg101, a cullin, RING-SH, and a
CC clathrin. The novel protein complex has virucide activity and can be used
CC to treat disorders as part of a vaccine. The protein complex and
CC composition are useful for detecting cells infected with a virus, for
CC identifying agents having antiviral activity, and for treating viral
CC disorders caused by retroviruses, rhabdoviruses, or filoviruses. This
CC sequence is a protein comprising the RING-SH complex of the invention.
XX
SQ Sequence 1048 AA;

Query Match 88.2%; Score 5483; DB 7; Length 1048;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPVQAQWTEFLSCPICYNEFDENVHKKPISLGCSTHTVCKTCLNKLHRKACPFDDQAINTD 60
DB 1 MPVQAQWTEFLSCPICYNEFDENVHKKPISLGCSTHTVCKTCLNKLHRKACPFDDQAINTD 60

QY 61 IDVLVFNFAQLQVGAQVDPHQSIKLSNLGENKHGEVAKKCVEDIALYLKPLSGGKGVAS 120
DB 61 IDVLVFNFAQLQVGAQVDPHQSIKLSNLGENKHGEVAKKCVEDIALYLKPLSGGKGVAS 120

QY 121 LNQSALSRPMQKLVTLVNCQLVEEGRVRAARSLGERTVTTELILQHNPQOOLSANL 180
DB 121 LNQSALSRPMQKLVTLVNCQLVEEGRVRAARSLGERTVTTELILQHNPQOOLSANL 180

QY 181 WAAVRARGCOFLGPAWQEBALKLVLLALEDDGSAISRKVLVLFVQRLPRFPQAKSTIG 240
DB 181 WAAVRARGCOFLGPAWQEBALKLVLLALEDDGSAISRKVLVLFVQRLPRFPQAKSTIG 240

QY 241 HVVQLLYRASCFCVTKRDESSLMQKEEPRSYEARLREHDAQIVHIAEAGLRISPEQW 300
DB 241 HVVQLLYRASCFCVTKRDESSLMQKEEPRSYEARLREHDAQIVHIAEAGLRISPEQW 300

QY 301 SLLYGDIAHKSQMSIIDKLOSPESFAKSVOELTIVLQRTGDPANLRLRPHLELLANI 360
DB 301 SLLYGDIAHKSQMSIIDKLOSPESFAKSVOELTIVLQRTGDPANLRLRPHLELLANI 360

QY 361 DPNDAVSPTEOLENAMVAVKTVVHGLVDFIQNSRKGHETPOPOPNKYKTSRCRDLR 420
DB 361 DPNDAVSPTEOLENAMVAVKTVVHGLVDFIQNSRKGHETPOPOPNKYKTSRCRDLR 420

QY 421 QOQCGPRGTCTCTFAHSQEBELKYLRLNKKINATVTRTFLPNKVGNNVTVTAGNIVSI 480
DB 421 QOQCGPRGTCTCTFAHSQEBELKYLRLNKKINATVTRTFLPNKVGNNVTVTAGNIVSI 480

QY 481 GSTETTGKIVPSTNGISNAENSVSQISRSRSTDLRALETVKVKGVKGANGAAGPSAD 540
DB 481 GSTETTGKIVPSTNGISNAENSVSQISRSRSTDLRALETVKVKGVKGANGAAGPSAD 540

QY 541 SVTENKIGSPKTPVSNVAATSAQPSNVGTSLNSVPQKSSPFLTRKVPVYPHSENIQYFQ 600
DB 541 SVTENKIGSPKTPVSNVAATSAQPSNVGTSLNSVPQKSSPFLTRKVPVYPHSENIQYFQ 600

QY 601 DPRTQIPFEVFPQYPTGYPPPTVPAGVAPCVPRFVRSNNVPESSLPASMPYADHYST 660
DB 601 DPRTQIPFEVFPQYPTGYPPPTVPAGVAPCVPRFVRSNNVPESSLPASMPYADHYST 660

QY 661 FSPDRMNSPYQPPPPQYGFVPPVPSGMVAPVYDSRRIRWPPMQRDDDIIRSNLPPM 720
DB 661 FSPDRMNSPYQPPPPQYGFVPPVPSGMVAPVYDSRRIRWPPMQRDDDIIRSNLPPM 720

QY 721 DMHSSVYQTSILRERYNSLDGYSVACOPPEPRTTVPDPREPCGHLKTSCEEQIRRKPD 780
DB 721 DMHSSVYQTSILRERYNSLDGYSVACOPPEPRTTVPDPREPCGHLKTSCEEQIRRKPD 780

QY 781 QWAQVHTQKAPLVSTLPAVQSTPFPSPFLSFVDPRADFSESVSGTKFEEHLSHSPWS 840
DB 781 QWAQVHTQKAPLVSTLPAVQSTPFPSPFLSFVDPRADFSESVSGTKFEEHLSHSPWS 840

QY 841 COTIGSCINAIDSEPKDVIANSNVLMOLDSDGVKRRVHLFETQRTTKEEDPIIFSDGP 900
DB 841 COTIGSCINAIDSEPKDVIANSNVLMOLDSDGVKRRVHLFETQRTTKEEDPIIFSDGP 900

QY 901 IISKWGAISRRTGYHTTDPVQATASQGSATKPIISVSDYVPYVNAVDSRWSSYGNAT 960
DB 901 IISKWGAISRRTGYHTTDPVQATASQGSATKPIISVSDYVPYVNAVDSRWSSYGNAT 960

QY 961 SAHYVERDRFVITDLSGHRKHSSTGDLISLELQQAQKSNLLOREANALAMQOKWNSLDE 1020
DB 961 SAHYVERDRFVITDLSGHRKHSSTGDLISLELQQAQKSNLLOREANALAMQOKWNSLDE 1020

QY 1021 GRHLTLNLSKEIELRNGELQSDYTEDA 1048
DB 1021 GRHLTLNLSKEIELRNGELQSDYTEDA 1048

RESULT 8
AAM40354
ID AAM40354 standard; protein; 869 AA.
XX
AC AAM40354;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 3499.
XX
KW Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.

XX OS Homo sapiens.
XX PN WO200153312-A1.
XX PD 26-JUL-2001.
XX PF 26-DEC-2000; 2000WO-US034263.
XX PR 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00682191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX PA (HYSE-) HYSEQ INC.
XX XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI: 2001-442253/47.
DR N-PSDB; AA159510.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX Example 6; SEQ ID NO 3499; 10078pp; English.
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA38442-AA42213) with neurotropic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX SQ Sequence 869 AA;
Query Match 71.1%; Score 4416; DB 4; Length 869;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 844; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 196 MQEALKLVLLALEDGSAISRKVLVLFVQRLPRFPQASKTSIGHVQQLLYRASCCKVT 255
DB 1 MQEALKLVLLALEDGSAISRKVLVLFVQRLPRFPQASKTSIGHVQQLLYRASCCKVT 60
QY 256 KRDEDSLMOLKEEPRSYEARREHDAQIVHIAEAGLRISPEQWSLLYGLDLAKSHMQ 315
DB 61 KRDEDSLMOLKEEPRSYEARREHDAQIVHIAEAGLRISPEQWSLLYGLDLAKSHMQ 120
QY 316 SIIDKLQSPESFAKSVOELTVLQRTGDPANLRLPHLELLANIDPNDAVPTWEQLE 375
DB 121 SIIDKLQSPESFAKSVOELTVLQRTGDPANLRLPHLELLANIDPNDAVPTWEQLE 180
QY 376 NAMVAVKTVHGLVDFONTYSRKGHETPOQPNKYKTSNCRDLROGGCGPGTCTFAH 435
DB 181 NAMVAVKTVHGLVDFONTYSRKGHETPOQPNKYKTSNCRDLROGGCGPGTCTFAH 240
QY 436 SQEELEKYLRLNKKINATVTFPLLNKVGNNVTVTAGNVISVIGSTETGKIIVPSTNG 495

DB 241 SQEELEKYLRLNKKINATVTFPLLNKVGNNVTVTAGNVISVIGSTETGKIIVPSTNG 300
QY 496 ISNAENSVSOLISRSSTDLRALETVKVKGVANGQNAAGPSADSVTENKIGSPKTPV 555
DB 301 ISNAENSVSOLISRSSTDLRALETVKVKGVANGQNAAGPSADSVTENKIGSPKTPV 360
QY 556 SNVAATSAGPSNVGTGLNSVQKSPFLTRVPVPPHSENIOYFODPRTQIIFEVPOYPQ 615
DB 361 SNVAATSAGPSNVGTGLNSVQKSPFLTRVPVPPHSENIOYFODPRTQIIFEVPOYPQ 420
QY 616 TGYYPPTVPAGVAPCVPRFVRSSNNVPESLPPASMPYADHYSTFSPRDRNNSPYQPP 675
DB 421 TGYYPPTVPAGVAPCVPRFVRSSNNVPESLPPASMPYADHYSTFSPRDRNNSPYQPP 480
QY 676 PPQYPGPVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRSNLPPMDVMHSSVYQTSLRER 735
DB 481 PPQYPGPVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRSNLPPMDVMHSSVYQTSLRER 540
QY 736 YNSLDGYYSVACQPPSEPRTTVPLPRPCGHLKTSCEQIRKPKDQWQAYHTQKAPLVSS 795
DB 541 YNSLDGYYSVACQPPSEPRTTVPLPRPCGHLKTSCEQIRKPKDQWQAYHTQKAPLVSS 600
QY 796 TLPVATQSTPPSPPLFSVDFRADPSESVSCTKFEEDHLSHYSPWSCGTIGSCINAIDSP 855
DB 601 TLPVATQSTPPSPPLFSVDFRADPSESVSCTKFEEDHLSHYSPWSCGTIGSCINAIDSP 660
QY 856 KDVIANSNAVLMDLSDGVKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRRTG 915
DB 661 KDVIANSNAVLMDLSDGVKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRRTG 720
QY 916 YHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSYGNATSAHVVERDRFVTDL 975
DB 721 YHTTDPVQATASQGSATKPIVSVDYVYVNAVDSRWSSYGNATSAHVVERDRFVTDL 780
QY 976 SGHRKHSSTGDLISLELOAKNSLLOREANALAMQKNSLDEGRHLTLNLSKEIEL 1035
DB 781 SGHRKHSSTGDLISLELOAKNSLLOREANALAMQKNSLDEGRHLTLNLSKEIEL 840
QY 1036 RNGLQ 1041
DB 841 RNGLQ 846
RESULT 9
AA40355
ID AA40355 standard; protein; 812 AA.
XX AA40355;
AC AA40355;
XX
DT 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 3500.
DE Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX Homo sapiens.
OS
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US034263.
PF
XX 23-DEC-1999; 99US-00471275.
PR
XX 21-JAN-2000; 2000US-00488725.
PR
XX 25-APR-2000; 2000US-00552317.
PR
XX 20-JUN-2000; 2000US-00598042.
PR
XX 19-JUL-2000; 2000US-00620312.
PR

PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI59511.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 6; SEQ ID NO 3500; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAI38642-AAI42213) with neurotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX Sequence 812 AA;

Query Match 65.4%; Score 4066.5; DB 4; Length 812;
Best Local Similarity 93.0%; Pred. No. 1.3e-302;
Matches 787; Conservative 2; Mismatches 0; Indels 57; Gaps 1;

QY 196 MOEALKLVLLALEDGSALSRLVLFVVQRLEPRFPQASKTSIGHVQLLYRASCFTVT 255
Db 1 MOEALKLVLLALEDGSALSRLVLFVVQRLEPRFPQASKTSIGHVQLLYRASCFTVT 60

QY 256 KRDESSLMLKEEPRSYEARREHDAQIVHIAEAGLRISPEQWSSLLYGLAHKSHMQ 315
Db 61 KRDESSLMLKEEPRSYEARREHDAQIVHIAEAGLRISPEQWSSLLYGLAHKSHMQ 120

QY 316 SIIDKLQSPESPAKSQVELTIVLQRTGDPANLRLRPHLELLANIDPNPDVSPTEQLE 375
Db 121 SIIDKLQSPESPAKSQVELTIVLQRTGDPANLRLRPHLELLANIDPNPDVSPTEQLE 180

QY 376 NAMAVKTVHGLVDFIQNSRKGHETPOQNSKYKTSKMRDLRQGGCPRGTCTFAH 435
Db 181 NAMAVKTVHGLVDFIQNSRKGHETPOQNSKYKTSKMRDLRQGGCPRGTCTFAH 240

QY 436 SQEELEKYLRLNKKINATVTRPPLNKGVNNTVTTAGNVISVIGSTETTKIVPSTNG 495
Db 241 SQEELEKYLRLNKKINATVTRPPLNKGVNNTVTTAGNVISVIGSTETTKIVPSTNG 300

QY 496 ISNAENSQILSRSTDSTLRALETVKVKVGKVGANGQNAAGSADSVTENKIGSPKTPV 555
Db 301 ISNAENSQILSRSTDSTLRALETVKVKVGKVGANGQNAAGSADSVTENKIGSPKTPV 360

QY 556 SNAVATSAGPSNVGTNELNSVPKQSPFLTRVPVPPHSENIQYFQDPRTQIFFEVPPQ 615
Db 361 SNAVATSAGPSNVGTNELNSVPKQSPFLTRVPVPPHSENIQYFQDPRTQIFFEVPPQ 420

QY 616 TGYIPPPPTVPAGVAPCPFRFVRSNNVPSSESLPPASMPYADHYSTFSPDRNNSPYOPP 675
Db 421 TGYIPPPPTVPAGVAPCPFRFVRSNNVPSSESLPPASMPYADHYSTFSPDRNNSPYOPP 480

QY 676 PPQYGPVPVPVPSGMYAPVYDSRRIRWPPMYQORDDIIRSNLPPMDVMHSSVYQTSLEER 735
Db 481 PPQYGPVPVPVPSGMYAPVYDSRRIRWPPMYQORDDIIRSNLPPMDVMHSSVYQTSLEER 540

QY 736 YNSLDGYYSVACQPPSPBRTTVPLEPRPCGHLKTSCEBQIRRKPDQWQAQYHTQKAPLVSS 795
Db 541 YNSLDGYYSVACQPPSPBRTTVPLEPRPCGHLKTSCEBQIRRKPDQWQAQYHTQKAPLVSS 566

QY 796 TLPVATQSPPTPSPPLFSVDPRADPSESUSGTFKFEEDHLSHYSPWSCGTIGSCINAIDSEP 855
Db 567 -----FSESUSGTFKFEEDHLSHYSPWSCGTIGSCINAIDSEP 603

QY 856 KDVIANSAVLMDLDSGDVKRRVHLFETQRTKEDPIIPFSDGPIISKWGAISRSSRTG 915
Db 604 KDVIANSAVLMDLDSGDVKRRVHLFETQRTKEDPIIPFSDGPIISKWGAISRSSRTG 663

QY 916 YHTTDPVQATASQGSATKPIISVSDYVPYVNAVDSRWSSYGNENATSSAHYVERDRPIVTDL 975
Db 664 YHTTDPVQATASQGSATKPIISVSDYVPYVNAVDSRWSSYGNENATSSAHYVERDRPIVTDL 723

QY 976 SGHRKHSSTGDLISLELOQAKSNLILLOREANALAMQOKWNSLDEGRHLTLNLLSKETEL 1035
Db 724 SGHRKHSSTGDLISLELOQAKSNLILLOREANALAMQOKWNSLDEGRHLTLNLLSKETEL 783

QY 1036 RNGLQ 1041
Db 784 RNGLQ 789

RESULT 10
ADC37580
ID ADC37580 standard; protein; 1125 AA.
XX AC ADC37580;
XX DT 18-DEC-2003 (first entry)
XX DE Human nucleic acid associated protein, NAAP-47.
XX KW Human; nucleic acid associated protein; NAAP; cytostatic;
KW antiarteriosclerotic; anticoagulant; neurotropic; neuroprotective;
KW cerebrotrophic; anti-HIV; antiallergic; antiinflammatory;
KW thymometric; gene therapy; cell proliferative disorder; cancer;
KW atherosclerosis; neurological disorder; epilepsy; Huntington's disease;
KW stroke; immune disorder; inflammatory disorder; AIDS; allergy;
KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection.
XX OS Homo sapiens.
XX PN WO2003046151-A2.
XX PD 05-JUN-2003.
XX PF 26-NOV-2002; 2002WO-US038445.
XX PR 27-NOV-2001; 2001US-0333925P.
PR 07-DEC-2001; 2001US-0340477P.
PR 14-DEC-2001; 2001US-0340362P.
PR 18-DEC-2001; 2001US-0342002P.
XX (INCY-) INCYTE GENOMICS INC.
XX BAUGHN MR, Becha SD, Bhatia U, Blake JJ, Burford N, Burrill JD;
PI Chang H, Chawla NK, Elliott VS, Emerling BM, Forsythe ID, Gandhi AR;
PI Gietzen KU, Gorvad AE, Griffin JA, Hafalia AJA, Jackson JL, Ho A;
PI Ison CH, Jackson AA, Jiang X, Jin P, Kable AE, Khare R, Lal PG;
PI Lee EA, Lee S, Lee SY, Li JK, Lu DAM, Ramkumar J, Richardson TW;
PI Sprague WW, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H;
PI Zheng W;
XX WPI; 2003-513642/48.
DR N-PSDB; ADC37640.
XX

PT New human nucleic acid associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
PT infections.

PS Claim 1; SEQ ID NO 47; 383pp; English.

XX The present invention relates to novel human nucleic acid associated
CC protein (NAAP) (I: ADC37534-ADC37593) and their coding sequences (J). The
CC NAAPs and their coding sequences are useful in diagnosing, treating and
CC preventing diseases or conditions associated with the decreased
CC expression or over expression of NAAP, such as cell proliferative (e.g.
CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
CC disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and
CC developmental (e.g. hypothyroidism, Cushing's syndrome) disorders, or
CC infections. These are also useful in assessing the effects of exogenous
CC compounds on the expression of nucleic acid and amino acid sequences of
CC NAAP.

XX Sequence 1125 AA;

Query Match 42.9%; Score 2665.5; DB 7; Length 1125;

Best Local Similarity 51.9%; Pred. No. 9.1e-195;

Matches 609; Conservative 141; Mismatches 295; Indels 129; Gaps 38;

QY 1 MPVQAQWTEFLSCPTCYNEFOENVHVKPISLGCSTVCKTCLNKLHRAKCPDQTAINTD 60

DB 1 MPVQAQWTEFLSCPTCYNEFOENVHVKPISLGCSTVCKTCLNKLHRAKCPDQTAINTD 60

QY 61 IDVLPVNFALLQVGAQVDPHOSIKL-SNLGENKHYEAKVCVEDIALYLKPLSGKGVA 119

DB 61 IDVLPVNFALLQVGAQVDPHOSIKL-SNLGENKHYEAKVCVEDIALYLKPLSGKGVA 119

QY 120 --SLNQSALSRMQRKLVNLVNCQVVEEGRVAMRAARSLGERTVTELLIHOHPQOLS 177

DB 120 --SLNQSALSRMQRKLVNLVNCQVVEEGRVAMRAARSLGERTVTELLIHOHPQOLS 177

QY 121 LNSTQSVLSRPMQRKLVNLVNCQVVEEGRVAMRAARSLGERTVTELLIHOHPQOLS 180

DB 121 LNSTQSVLSRPMQRKLVNLVNCQVVEEGRVAMRAARSLGERTVTELLIHOHPQOLS 180

QY 178 ANLMAAVRARGCOFLGPAQOEALKLVLLALEDGSALSRLKLVLFVQVLEFRFPQASKT 237

DB 178 ANLMAAVRARGCOFLGPAQOEALKLVLLALEDGSALSRLKLVLFVQVLEFRFPQASKT 237

QY 181 SNLMAAVRARGCOFLGPAQOEALKLVLLALEDGSALSRLKLVLFVQVLEFRFPQASKT 240

DB 181 SNLMAAVRARGCOFLGPAQOEALKLVLLALEDGSALSRLKLVLFVQVLEFRFPQASKT 240

QY 238 SIGHVVQLLYRASCFTVKTRDESSLQMLKEEPSYEAALRRHDAQIVHAMEAGLRISP 297

DB 238 SIGHVVQLLYRASCFTVKTRDESSLQMLKEEPSYEAALRRHDAQIVHAMEAGLRISP 297

QY 298 EQWSSLLYGLAHKSHMQSIIDKQSPESPAKSVQELTVLQRTGDPANLRLRPHLELL 357

DB 298 EQWSSLLYGLAHKSHMQSIIDKQSPESPAKSVQELTVLQRTGDPANLRLRPHLELL 357

QY 301 DQWSSLLYGDQSHKSHMQSIIDKQSPESPAKSVQELTVLQRTGDPANLRLRPHLELL 360

DB 301 DQWSSLLYGDQSHKSHMQSIIDKQSPESPAKSVQELTVLQRTGDPANLRLRPHLELL 360

QY 358 ANIDPNPDVAVPTWOLENAMVAVTVVHGLVDFIONYSRKGHETPQOPNSKYKTMCR 417

DB 358 ANIDPNPDVAVPTWOLENAMVAVTVVHGLVDFIONYSRKGHETPQOPNSKYKTMCR 417

QY 361 ANIDPSDPAPPPWEQENGLAVRIVVHGLVDYIQNHKKGADQQQPPQSHKYKTYMCR 420

DB 361 ANIDPSDPAPPPWEQENGLAVRIVVHGLVDYIQNHKKGADQQQPPQSHKYKTYMCR 420

QY 418 DLKQCGCGPRGTCTCTFAHSELEKYLRLNKKINATVTFPLINKVGNNVTVTAGNVI 477

DB 418 DLKQCGCGPRGTCTCTFAHSELEKYLRLNKKINATVTFPLINKVGNNVTVTAGNVI 477

QY 478 SVTGSTTETTKIVPS--TNGISNAENSVSQILSRSTSTLRALETVKVKGVKGVANGQNAAG 536

DB 478 SVTGSTTETTKIVPS--TNGISNAENSVSQILSRSTSTLRALETVKVKGVKGVANGQNAAG 536

QY 478 PDEGAVDLPSRKPALPANGIVSTGNTVTLIPRGTDG---SYDSSLKPGKI----- 525

DB 478 PDEGAVDLPSRKPALPANGIVSTGNTVTLIPRGTDG---SYDSSLKPGKI----- 525

QY 537 PSADSVTENKIGSPKTPVSNVAATAGPSNVCTELNSVPQKSSPFLTRVPV----- 588

DB 537 PSADSVTENKIGSPKTPVSNVAATAGPSNVCTELNSVPQKSSPFLTRVPV----- 588

QY 526 ---DHLSSAGSPFPPOLLESVPKSIASLPVNP-----HSIPFRGADLPMPVTKPLQMPV 578

DB 526 ---DHLSSAGSPFPPOLLESVPKSIASLPVNP-----HSIPFRGADLPMPVTKPLQMPV 578

QY 589 -----YPPHSENTQYFQDPR-TQIPFEVPOYPTQYPPPPPTVPAGVAPCVPRFVRNNV 642

DB 589 -----YPPHSENTQYFQDPR-TQIPFEVPOYPTQYPPPPPTVPAGVAPCVPRFVRNNV 642

QY 579 RGSQQLPAQOTDV-YVQDPRGAAPPPEPAPYQGMYYTTPP-----QCVSRFVR--- 626

DB 579 RGSQQLPAQOTDV-YVQDPRGAAPPPEPAPYQGMYYTTPP-----QCVSRFVR--- 626

QY 643 PESSLPASMPYADHYSTFSPRDMNSSPYQPPPPPPQYGVPPVPSGMVAPVYDSRIWR 702

DB 643 PESSLPASMPYADHYSTFSPRDMNSSPYQPPPPPPQYGVPPVPSGMVAPVYDSRIWR 702

QY 627 PPSAPAPAPPYLDHYPPYL-QERVVNSQYGTQFQQ-----YPPI-----YPSHYDGRVYP 677

DB 627 PPSAPAPAPPYLDHYPPYL-QERVVNSQYGTQFQQ-----YPPI-----YPSHYDGRVYP 677

QY 703 PPMYQRDIIIRSNLUPPMDVMHSSV--YQTSURERYNSLDGYYSVACQP-----PS-----E 752

DB 678 APSYTREIFRESPI-PIEIPPAAPVPSRSEYQIESTYYPVAPHPTQIRPSYLRFP 736

QY 753 PRTTVPLPREPCGHLKTSCEQIRKPDQWQYHTQKAPLVSSLTPLVATQSTPTPSPPLFS 812

DB 737 PYSRLUPPPPPQP-----HPSLDELHRRKRIIMQAQLEBK---VISPPPPFA-PSPTLP-PTPH 787

QY 813 VDFRADFSESYSCTKFEEDHLSHSPWCGTIGSCINAIIDSEPKDVIANSNVLDLDSG 872

DB 788 PEEFLDEDLKVAG-KYKNDYISQYSPWSCDTIGYIGTKDAKPKDVVAAGSVEMNVESK 846

QY 873 DVKRRVHLFEQRTTKB--EDPIIFSDGPIISKWGAISRSRTGYHTTDPVQATASQGS 930

DB 847 GM--RDQRLDLQRRAAETSDDLIPFGDRPTVRFCAISRTSKTIYQAGPMQAMAPOGA 904

QY 931 ATKPISVDYYPYVNAVDSRW--SSYG--NEATSSAHYVERDRFIVTDLGHRK--HSST 984

DB 905 PTKSINISDYSY--GTHGGMGASPYSPHONIASOGHFSERERISMSEVASHGKPLPSAE 962

QY 985 GDLSLELEQQ-----AKSNSLL-LQREANALAMQO-----KWSLDEGRHLLTLNL-- 1028

DB 963 REQLRELEQLNLHQSQQTLQGLEREANTLAGOSQPPPPPPKPMGMISSSEQLSLELHQ 1022

QY 1029 ISKEIELANGELQSDYTDATDKPDRIELESALDTE--PDGQSEPIEILDQLGI 1086

DB 1023 VEREIGKRTRELSME-NQCSLDMK-----SKLNTSKQAENGQPEQPNKVPABDLTL 1072

QY 1087 SSCQDQLNGMAVENGHVPVQHQKPEPPKQKQSL 1120

DB 1073 TFFSSD-----VPNGSALTQENISLLSNKTSLSL 1099

RESULT 11

ADF5448
ID ADF5448 standard; protein; 1109 AA.

XX ADF5448;

XX 12-FEB-2004 (first entry)

XX Human novel polypeptide #12.

XX human; brain disease; mental disorder.

XX Homo sapiens.

XX JP2003245081-A.

XX 02-SEP-2003.

XX 25-FEB-2002; 2002JP-00047501.

XX 25-FEB-2002; 2002JP-00047501.

XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.

XX WPI; 2003-857164/80.

XX N-PSDB; ADF55399.

XX New DNA derived from KG-1 cDNA library and encoded polypeptide, useful for treating mental disorders.

XX Claim 4; SEQ ID NO 12; 401pp; Japanese.

XX The invention relates to a DNA which encodes a novel polypeptide. A vector containing the DNA is useful as a reagent in estimation of standard substance. The antibody is useful for detecting the vector containing the DNA and for screening substances and compounds that interact specifically with the vector containing the DNA. The vector containing the DNA is used as a pharmaceutical, in treatment or as preventive agent with respect to disease e.g., brain diseases preferably mental disorders. The present sequence represents the amino acid sequence

CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention
XX
SQ Sequence 421 AA;

Query Match 33.9%; Score 2106; DB 4; Length 421;
Best Local Similarity 99.2%; Pred. No. 1.7e-152;
Matches 395; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 644 ESSLPPASMPYADHYSTFSPDRMNSPYQPPPPQYGPVPPVPSGMYAPVDSRRIRP 703
DB 1 ESSLPPGSMYPADHYSTFSPDRMNSPYQPPPPQYGPVPPVPSGMYAPVDSRRIRP 60

QY 704 PMYQRDDIIRNSLPPMDVHSSVYQTSLSRERNVSLDGYYSVACQPPSPRTTVPILPREP 763
DB 61 PMYQRDDIIRNSLPPMDVHSSVYQTSLSRERNVSLDGYYSVACQPPSPRTTVPILPREP 120

QY 764 CGHLKTSCEQIRKPDQWAQYHTQKAPLVSSSTLPVATQSPPTPPLFSVDFRADFSESV 823
DB 121 CGHLKTSCEQIRKPDQWAQYHTQKAPLVSSSTLPVATQSPPTPPLFSVDFRADFSESV 180

QY 824 SGTKFEDHLSHYSPWSCGTIGSCINAIDSEPKDVTANSNAVLMDLDSGDKRRVHLFET 883
DB 181 SGTKFEDHLSHYSPWSCGTIGSCINAIDSEPKDVTANSNAVLMDLDSGDKRRVHLFET 240

QY 884 ORTKEDDIIIPSDGPIISKWGAISRSTGYHTTDPVQATASQGSATKPIVSVDYVPY 943
DB 241 ORTKEDDIIIPSDGPIISKWGAISRSTGYHTTDPVQATASQGSATKPIVSVDYVPY 300

QY 944 VNAVDSRWSSYGNATSSAHYVERDFIVTDLGSKRKHSTGDLISLELQQAQKSNLLQ 1003
DB 301 VNAVDSRWSSYGNATSSAHYVERDFIVTDLGSKRKHSTGDLISLELQQAQKSNLLQ 360

QY 1004 REANALAMQKWNLSLDEGRHLTLNLLSKIELRNGELQ 1041
DB 361 REANALAMQKWNLSLDEGRHLTLNLLSKIELRNGEVK 398

RESULT 13
AAM42140
ID AAM42140 standard; protein; 306 AA.
XX
AC AAM42140;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 7071.
XX
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX

PD 26-JUL-2001.
XX
26-DEC-2000; 2000WO-US034263.
XX
23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-0052317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.

(HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
XX Zhou P, Goodrich R, Drmanac RT;
DR WPI; 2001-442253/47.
XX N-PSDB; AAI61296.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 2; SEQ ID NO 7071; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nontropic, and the
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
SQ Sequence 306 AA;

Query Match 25.6%; Score 1593; DB 4; Length 306;
Best Local Similarity 98.7%; Pred. No. 2.6e-113;
Matches 293; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 514 TLRALETVKVKGVKGANGQNAAGPSADSVTENKIGSPKPTPVSNVAATSAGPSNVGTEN 573
DB 4 TLRALETAKVKGVKGANGQNAAGPSADSVTENKIGSPKPTPVSNVAATSAGPSNVGTEN 63

QY 574 SVPQKSSPFLTRVPVYPHSENIQYFQDPRTOIPFEVPOYPQTGYPPPTVPAGVAPCV 633
DB 64 SVPQKSSPFLTRVPVYPHSENIQYFQDPRTOIPFEVPOYPQTGYPPPTVPAGVAPCV 123

QY 634 PRFVRSNNVPESLPPASMPYADHYSTFSPDRMNSPYQPPPPQYGPVPPVPSGMYAP 693
DB 124 PRFVRSNNVPESLPPASMPYADHYSTFSPDRMNSPYQPPPPQYGPVPPVPSGMYAP 183

QY 694 VYDSRRIRWPMYQRDDIIRNSLPPMDVHSSVYQTSLSRERNVSLDGYYSVACQPPSEP 753
DB 184 VYDSRRIRWPMYQRDDIIRNSLPPMDVHSSVYQTSLSRERNVSLDGYYSVACQPPSEP 243

QY 754 RTTVPLPREPCGHLKTSCEQIRKPDQWAQYHTQKAPLVSSSTLPVATQSPPTPSP 810
DB 244 RTTVPLPREPCGHLKTSCEQIRKPDQWAQYHTQKAPLVSSSTLPVATQSPPTPSTL 300

RESULT 14
AAM42141

Db 1 MPIQAFSWTDFLNCPICCNEFAASQRCVPSVLCGCHTICKLCLTTLNRCQCFDQTVIVSD 60
 Qy 61 IDVLVNEFALLQV-----CAQVDPHQSIKLSNIGENKXVYVAKKVEDLALYLKPL 112
 Db 61 IDNLPINHALQVLDKSELELAPPPSVQKIEPEHL---KYQLGQRCIEELALHLKSF 117
 Qy 113 SGGKGVASINQSALSRLPMQKLVTLVNCQVBEGRVRAARAARSIGERTVTTELILQHON 172
 Db 118 LNLNG---NGNLLTRPMLRKLVTLVNCLMEEGRVRAARAARSIGERTVTTELILQHON 173
 Qy 173 POOLSNLWAAVRARCCQFLGAMQBEALKVLVLALEDGALSARKVLVLFVQRLSPRP 232
 Db 174 POOLSNLWAAVTRCQFLGAMQBEVLKVLVLALEEGSALSARKVLVNFVQRLSPRP 233
 Qy 233 QASKTSIGHVVOLLYRASCFTKTRDESSLMQKEEFYSALREHDAQIVHIAEAG 292
 Db 234 QASKTSIGHVVOLLYRASCFTKREADSLSLMQKEEFRTYDALREHDAQIVQIATEAG 293
 Qy 293 LRISPQWSSLLYGLDLAHKSHMQSIIDKLQSPESFAKSVOELTIVLQRTGDPANLRLRP 352
 Db 294 LRISPQWSSLLYGLVHSHMQSIIDKLQTPSSFAQSVOELVIALQRTSDPAKLSLHH 353
 Qy 353 HLELLANIDPNDAVPTWEQLENAVMVKTVVHGLVDFIQNYS-RK-----GHETPQ 404
 Db 354 HLKYLANIDPCAB-VAP-NSVLAEALDAVRHSVVGVLVNFQHHGVKQAQDGISGGSGGT 411
 Qy 405 PQNSKYKTSMDRLRQCGCRGTNCTFAHSGOELEKYLRLNKKINATVTFPLLNKVG 464
 Db 412 TNSNPKYKISLCRDLNVRCVRGSSCTFAHSGOEVEERYARNR----- 455
 Qy 465 VNNVTVTAGNVISVIGSTETGKIVPSTNGISNAENSVSQIISRSTDTLRALETVKV 524
 Db 456 ----- 455
 Qy 525 GKVGANGQNAAGSADSVTENKIGSPKTPVGNVAATSAGPNSVGTSLNSVP--QKSSPF 582
 Db 456 -----GKHKTPLA-----LOGPPAVGVGAIKKPLGEQEGPP 487
 Qy 583 LTRVPVYPHSENIQYQDPRTQIPPEVPQYQOTGYPPPTVPAGVAPCVPRFVRSNNV 642
 Db 488 LGNMPMLPMSP-MHYMGSPR-----GYLDP-----SLGLSP-----GGGL 522
 Qy 643 PESSLPAS---MP--YADHYSTF---SPDRMNSSPYQPPPPQYGPVPPVPSGMYAPV 694
 Db 523 PPSHHSPIITRLIVPSRYDSRPSGFGGCTPR-----IPSPREYOANPAP----- 566
 Qy 695 YDSRRIRPPMYQRDDIIRSNLSLPPM---DVHSSVYQTSRLRERNSLDGYYSVACQPP- 750
 Db 567 --TORNANPNFNSVNSNLHKGMYLPASGGDVFLA-----NPWEQAVLAQQQHP 614
 Qy 751 -----SEPRTTVPLPREPCGHLKTSCEBQ-IRKPDQWQYHTQKAPLVS-STLPVATOS 803
 Db 615 QHPQOQQPPSSKNPSRPLSILPATADTSFFFEKKPPNSVSIDLDRVPEVNVDAVPLFRS 674
 Qy 804 PTPPSPFLSVDFFADSESVSGTKFEEDHLSHYSPWSCGTIGSCINAIIDSEPKDVIANSN 863
 Db 675 -----NNHNNNNNNHNNNNHSSLLFWN-NTGKDSANFVS-----DSLDDDD 720
 Qy 864 AVLMDLDSGDVKKRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSRTG 915
 Db 721 ASTFDVPTGSSMLSIY-----GPICPK-----SSTTG 747

Search completed: August 4, 2005, 01:36:14
 Job time : 185 secs

This Page Blank (Cont.)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:26:33 ; Search time 26 Seconds
(without alignments)
4407.468 Million cell updates/sec

Title: US-10-619-992-2
Perfect score: 6215
Sequence: 1 MPVQAAQWTEFLSCPICYNE.....MSEDKNDFLKPANGKVVNS 1191

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	843	13.6	1048	2 T23764	hypothetical prote
2	219.5	3.5	1186	2 T13050	hypothetical prote
3	207.5	3.3	1198	2 T49726	hypothetical prote
4	206	3.3	1611	2 T38236	hypothetical prote
5	203	3.3	786	2 T01456	extensin homolog F
6	202	3.3	760	2 T06291	extensin homolog T
7	194.5	3.1	1279	2 T18312	hypothetical prote
8	193.5	3.1	744	2 E86255	hypothetical prote
9	192.5	3.1	944	2 T28734	hypothetical prote
10	191	3.1	368	2 C29356	hydroxyproline-ric
11	191	3.1	2160	2 T20241	hypothetical prote
12	190.5	3.1	760	2 F86387	probable Pco kinase
13	190.5	3.1	842	2 S60402	protein kinase CLA
14	188.5	3.0	1819	2 T32008	hypothetical prote
15	183	2.9	280	2 T03236	extensin precursor
16	181.5	2.9	530	2 T48217	hypothetical prote
17	181.5	2.9	2282	2 T42717	DNA-binding protei
18	181	2.9	1257	2 T00486	serine/threonine-s
19	179	2.9	2562	2 T14266	Xin protein - chic
20	179	2.9	2783	1 A41948	alpha-retoprotein
21	178	2.9	1013	2 T33470	hypothetical prote
22	177.5	2.9	416	2 JU0465	extensin precursor
23	177.5	2.9	620	2 S06733	hydroxyproline-ric
24	177.5	2.9	1375	2 S48375	hypothetical prote
25	177	2.8	1812	2 I49350	breast/ovarian can
26	175.5	2.8	429	2 T06296	extensin-like prot
27	174.5	2.8	951	2 T47617	extensin-like prot
28	173	2.8	1262	2 T25168	hypothetical prote
29	173	2.8	1794	2 T38459	hypothetical diver

ALIGNMENTS

RESULT 1

T23764
hypothetical protein M142.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23764, T27112
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19795
A:Accession: T23764
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1048 <W14>
A:Cross-references: UNIPROT:O45962; EMBL:Z73428; PIDN:CAA97810.1; GSPDB:GN00021; CESP:M142.6
A:Experimental source: clone M142
R:Steward, C.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z20313
A:Accession: T27112
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1048 <W12>
A:Cross-references: EMBL:Z99276; PIDN:CAB16481.1; GSPDB:GN00021; CESP:M142.6
A:Experimental source: clone Y52D3
C:Genetics:
A:Gene: CESP:M142.6
A:Map position: 3
A:Introns: 21/3; 122/3; 381/3; 579/3; 647/2; 683/1; 783/2; 855/1; 871/2; 909/1; 993/1

Query Match 13.6%; Score 843; DB 2; Length 1048;

Best Local Similarity 30.6%; Pred. No. 9.8e-39;
Matches 238; Conservative 99; Mismatches 278; Indels 162; Gaps 23;

QY	4	QAAQWTEFLSCPICYNEFDENVHVKIPISLGCSHTVCKTCLNKLHR--KACPFQDTAINTDI	61
Db	6	QGGQWQEVLCSSICNRHENE-TFLPVSLICGHVICRKAEPENQTKPCPHDDWKTTHSP	64
QY	62	DVLPNPFALLQVGAQVDPHQSIKLSNLGNKHCHYEYAKKCVEDLALYLPKSGKGVASL	121
Db	65	SEYPNNVALLSVI---FPRKQCMTLGSAVSEAKRVDQLSIQ-IAKFFREADSERG-GTV	119
QY	122	NQSALSRPMOKLVTLVNCQVVEEGRVRAAARSLSGERTVTTELQHNQPOOLSANLW	181
Db	120	SRESRLTQKRVALLCYQWEVVGRLTKLQKCGISERVMIETLSQSNTHVSSQLW	179
QY	182	AAVRARGCQFLGPAQOEALKLVLLEEDGSALSRLVLFVVQRLERPRPQAKTSIGH	241
Db	180	SAVRARGCQFLGPAQODDLRLILMTLETGECIARKNLVMYVVQTLASDYPQVSKTCVGH	239
QY	242	VVQLYRASCFCVTKTRDESSLMQLKEEPRSRYSALRREHDAQIVHVTAMAGLRISPEQWS	301
Db	240	VVQLLYRASCFLNLRKDGESSLMQLKEEPRTRYESLRREHDSQIVQIAFESGLRIGPDQWS	299

A:Experimental source: strain Bristol N2; clone F26G5
C:Genetics:
A:Gene: CESP:F26G5.9
A:Map position: 5
A:Introns: 12/3; 48/3; 258/3; 327/3; 643/1; 916/2

```
Query Match          3.1%; Score 192.5; DB 2; Length 944;
Best Local Similarity 17.7%; Pred. No. 0.0096;
Matches 215; Conservative 146; Mismatches 410; Indels 445; Gaps 56;

QY 7 QWTEFLSPICYNEEDENVHKPISLGCSTHTVCTCLNKLHKACPFQDTAINTDIDLVPV 66
DB 14 QWAVVECPICNYID----KPMQMGCGHTLSTCIGRL-----VDQVKT 54
QY 67 NFALLQLVGAQVPDQHSIKLSNGENKHYEVAKKVEDLALYLKPLSGKGVASLNQSA- 125
DB 55 NNA-----EELGRPR-----GPLLHPDDIGPGFGLGGFAD 85
QY 126 -LSRPMQRLVTLVNCQLVEEGRVRAEARSIGERTVTLELLOHNPQQLSANLWAAV 184
DB 86 QEDQHFHIDLPRDPRMPNDPNWVGI-----VVLPRGPE----- 122
QY 185 RARGQQLGPAMQEBALKLVLLALEDDGALSRLVLFVYVQRLPEPRFQASKTSIHVVQ 244
DB 123 -ARATEIKCPECKPT-----LVPADGLPVNRYVQE--IVQVAPLFXDR-----HLVK 168
QY 245 LLYRASCVKYTRDESSLMQKESFRSEALRREHDAQIVHAEAGLRISPEOWSSLL 304
DB 169 LCNQ--CEAVLSQGVYDFCQCEETGRK-----ICSTCAIRL----- 203
QY 305 YGDLAHSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRPHLELLANIDNP 364
DB 204 -----HNGH-----QVKEKALTSDVREMKQKISDASH-----AAFQLENLKLRL 245
QY 365 DAVSPTWE--QLENAMVAVKTVVHGLVDFIQNYSRKGHETPOPNQSKYKTSMCRLRQQ 422
DB 246 QSVGGTIEAKALDKLTSIKIFEFMLSTF-----DSKIK----- 279
QY 423 GCGPRTGNTFAHSEBELEKYLRLNKKINATVTRTPLLNKGVNNTVTYTAGNVISVGS 482
DB 280 -----ENSTMDLMBEVKK-----AEKISKVY----- 301
QY 483 TETTKGIVPSTNGISNAENS-----VSQLSRSTDSLRLALETVKVKVGKVGANGQNA 535
DB 302 SEAGRGVDVMLTAIESAFNSYNDPPEKLQEMGFQID-----EPTGSVNQI----- 347
QY 536 GPSADSVTENKIGSPKPTVSNVAATSAGPSNVGTSLNSVPQKSSPFLTRVPVYPHSEN 595
DB 348 -PASEGGAER-----VNEVAAPAAG-----RPIPNARNLRGALNHAFAQVARN 389
QY 596 IQYFQDPRTOIPFEVQYQOTGYVPPPPVPVAGVAPCVFRVRSNNVPSSSLPPASMPYA 655
DB 390 -----RNRQARGVQOQDLPLPLPPPP-----HQLVMPPPAMPPP 424
QY 656 DHYST-----FSR-DMNASSPQP-----PPQYGPVPPVPSGMYAPV 694
DB 425 QHMQAORLWGGFQFQRMN-----QPMMAQVGGPRGQVIMGHPPPMHMQOOGPRMMY--M 478
QY 695 YDSRRIWRPMPYQORDIIRSNLPPMDVMHSSVYQTSLRERNYNSLDGYYSVACQPPSEPR 754
DB 479 HQQHH-----PMHQVED-----QMQRQHORMQPNMN--NIMHGNMQIAFQQHMNN 523
QY 755 TTVPILPREPCGHLKTSCEQIARKPDQWAOY-----HTQKAPLVSTSLPVA 800
DB 524 ANVPMPPQPAHPQ---QQQQQRHNEGQFQGGPPPPPPQRIQHQOQMOQYHQQLQVQ 580
QY 801 TQSPFP-----SPLFSVDFA-----DF-----SEVSQTKFEEDHLSHYSP-- 838
DB 581 QOQQQOPRAQPRFRPTQQOQQOQLQIDLQAMNDRQFQQLADEIVAEQMVQNAVDAAPAD 640
QY 839 -----W-----SCGTIG-----SCINAIIDSEPKQV---IANSNAVIMDLSDGVKR 876
DB 641 IPRADWEELQVQRGAGDGFINFHPPEPIVEEVEEQPEVDQVQVNVGQVVENVRAGDINN 700
```

```
QY 877 RVHLFETQRTKEBDPII--PFSQDPIISKWGAISRSSRTGYHTTDPVQATASQGSATKP 934
DB 701 VIHF-----EEVPVLEPFDVGVNDHNGIENNVVEEQEQPDD-----VP 740
QY 935 ISVSDYVPYVNAVDSRWSSYGEATSSAHYVERDRFIVTDLSGHRKHSSTGDLALSLEQQ 994
DB 741 PQINNFEV-----DDMREQVDAIEDEEF-----EPVLHR-----AAIEFOE 777
QY 995 AKNSNLLQ--REANALAMQOKWNSLDEG--RHILTLLLSKEIELRNGELQSDYTEDATD 1050
DB 778 AADQDALBQDYRERNLLDHEAENEEDDDGLPYEEDDEDEEVHVRNSVVR--RVQRPD 836
QY 1051 TKPDRDIELELSAL-----DDEPDGQSEPIEILDIOLGISSONDQLLNGMAVENG 1102
DB 837 AAHISEVTRRSARLARNSNPTDEPSTSS-----GAAGPSN 874
QY 1103 HPVQOHOKEPPKQKKQ 1118
DB 875 RP--PHSPTPPPGNKR 888

RESULT 10
C29356
hydroxyproline-rich glycoprotein (clone Hyp2.13) - kidney bean (fragment)
C:Species: Phaseolus vulgaris (kidney bean)
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: C29356
Mol. Cell. Biol. 7, 4337-4344, 1987
R:Corbin, D.R.; Sauer, N.; Lamb, C.J.
A:Title: Differential regulation of a hydroxyproline-rich glycoprotein gene family in wo
A:Reference number: A29356; MUID:88142825; PMID:3437892
A:Accession: C29356
A:Molecule type: mRNA
A:Residues: 1-368 <OR>
A:Cross-references: UNIPROT:Q09085; EMBL:M18095; NID:g169348; PIDN:AAA33765.1; PID:g1693
A:Experimental source: cv. Klevitsboon Koekoek
C:Superfamily: hydroxyproline-rich glycoprotein
C:Keywords: glycoprotein

Query Match          3.1%; Score 191; DB 2; Length 368;
Best Local Similarity 24.3%; Pred. No. 0.0028;
Matches 93; Conservative 30; Mismatches 153; Indels 106; Gaps 16;

QY 468 TTTT--AGNVISVIGSTETTGKIVPSTNGI-----SNAENSVSQL 506
DB 3 TVEVTEVGGKTIKSYGNTKNGKYSITVEGLDYVYKGTVCQAQLHAPPKSGRCSIPTK 62
QY 507 ISRSSTDSTLRALETVKVKVGKVGANGQNAAGPSADSVTENKIGSPKPTVPS-NVAATSAGP 565
DB 63 LNEGTKLALKSKDYEVVLK-----AKPAYA-----PKPYDCEKHKHSPTP 105
QY 566 SNVGTSLNSVP---QKSSPFLTRVPVYPHSENIQYFQDPRTOIPFEVQYQOTGYVPP 621
DB 106 YHKPYVYNSPPPPYVYKSPPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 163
QY 622 PPTVAGVAPCVFRVRSNNVPSSSLPPASMPYADHYSTFSPDRDMNSSPY---QPPPPQ 678
DB 164 PPPSP---SPPPPPYVYKSPPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 215
QY 679 PVGPVP---PVPSGMVAPVYDSRRIRWRPMVQORDDIIRSNLPPMDVMHSSVYQTSLR 733
DB 216 FSPPPPPYVYKSPPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 264
QY 734 ERYNSLDGYYSVACQPPSEPRTTVPL-----PREPCGHLKTSCEQIARKPDQWAOYHT 787
DB 265 -----PPPDPSPPPPYVYKSPPPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 296
QY 788 QKAPLVSTSLPVATQSPPTPSP 809
DB 297 PPPPSPPPPPPYVYKSPPPSPSP 318
```

RESULT 11
T20241
hypothetical protein H05L14.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20241; T23049
R:Wilkinson, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19241
A:Accession: T20241
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2160 <WIL>
A:Cross-references: UNIPROT:O17709; EMBL:Z75533; PIDN:CAA99823.1; GSPDB:GN00019; CESP:HC
A:Experimental source: clone C5464
R:Barlow, K.
submitted to the EMBL Data Library, October 1997
A:Reference number: Z19662
A:Accession: T23049
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2160 <WIL>
A:Cross-references: EMBL:Z99772; PIDN:CAB16922.1; GSPDB:GN00019; CESP:H05L14.2
A:Experimental source: clone H05L14
C:Genetics:
A:Map position: 1
A:Introns: 49/3; 130/3; 190/3; 237/3; 290/1; 463/3; 507/1; 524/2; 545/1; 636/2; 663/2; 7
Query Match 3.1%; Score 191; DB 2; Length 2160;
Best Local Similarity 18.9%; Pred. No. 0.041;
Matches 233; Conservative 162; Mismatches 447; Indels 394; Gaps 59;
QY 45 LHRKACPFQTAINTDIDL-----PVNFALLQLVCAQ----VPDQHSIK 85
DB 494 VHFQICP-----PTNEDLHLISRLKLNLRKQYSGNFGSTFSSGAERKGNIPDKLLDS 549
QY 86 LSNLG----ENKHY----EVAKKCVEDLALYKPLSGGKGVASLNQSAISRPMORKLVTL-- 137
DB 550 PSNMGSDSESNDFPENSEQC-----SVSSADKGISSNSQT--SDVPKGVITLDP 599
QY 138 -----VNCQVVEEGRVAMRAARSLGRTVTTELILQHONPOOLSANLWAAVRARGCOF 191
DB 600 GTVCVSLTVKISEE-----SHKLG-----ELEAKKENDQI----- 631
QY 192 LGPAMQBEALKVLLALEDGSAISRKVLVFFVQRLEP-----RFPQASKTSIGHVV 243
DB 632 -----LEPSTSVLFERTAPPPVDSRPTSRFPQSKSN----- 664
QY 244 QLLYRASCPRKTKRDEBDSLWOLKEFRSYEARLREHDAQIVHIAEAGLRISPEQWSSL 303
DB 665 -LITRLNVNEI--HSLKSLMDLLKQEKMSQEQDDSDVD-----DESSVESSESPTSS 717
QY 304 LYGDLAHSHMQSIIDKLOSPESFAKSVQELTIVLQR--TGDPA-----NLNLRPHLEL 356
DB 718 ITISVASPTDPEAKRQKKEKKARQLSKQEQIKQKRLERAEEAAKQRELNLAEHKKRLAK 777
QY 357 LANIDNPDAV-----SPTWEQLENAMVAKTVVHGLVDFFQNYSRKGHETP 403
DB 778 YAALVAKPTVTKTDEQPNKKSTQLQPRVEKTEKGNISVSVSSDRQVE-----NRTG---- 828
QY 404 QPQNSKYKTSMCRDLRQCGGCRGTNC--TPAHSQEELEKYRLRNKKINATVTFPLLN 461
DB 829 -----SFPSPSTRLSNSNIP-----SCFATPDSQPSQLQSFVNSNDLDA----- 870
QY 462 KVGNNVTVTAGNVISVIGSTETTKIVPSTNGISNAENSVSQLISRSTDLRALETV 521
DB 871 DTSVNLISITTDGVHLMVCVDKA-AQGVVVEATH-VGESEN-----TIEVVEDA 916
QY 522 KKVGVKGVANGQNAAGSADSVTENKIGSPPKTPVSNVAATSAG--PSNVGTETLNSVPQKS 579
DB 917 EKSFEILVCCQRR-----KVT-----PFTPRRKQDSSSGDEQPAESSTSAFDIPEHP 963

QY 580 S---PFLTRVVPVPPHSENIQYFQPRTRQIPEVPOY----POTGYPPPPPTYPAGVAPCV 633
DB 964 PVIVFPRNAPIVYSSTQNI-----PQIHPLPFRFLAMPQFSGHPPPPVIMG----- 1012
QY 634 PRFVSNVNPESLPPASMPYADHYSTFSPRDMNSSPVQPPPPPOPYGVVP--PV----- 686
DB 1013 PPIIRPFQAP---PPQFIPOSNTG-----RPPQFFPIGVVPIPIQLPQW 1055
QY 687 ----PSGMYA--PVYDSRRIRPPMYQ-----RDD-----ITRSN-----SLP 718
DB 1056 RMSGFPVMMRGFPVYCQVPVPMHPMQNVQNMSSSRQDFGNVRLIRPNQKPLEITAP 1115
QY 719 PMDVHSSVYQ-----TSRERYNSLDGYYSVACQPSSEPRTV 757
DB 1116 PVSASNSKIEKMGKVRPEPRRKEIINVSRSQVRNKGKLDQF-----KTQKOTV 1167
QY 758 PLPREPCGHLKTSCEQIRRRKPDQWAOYHTQKAPLVSTLPVATOSPTPPSLFSVDFRA 817
DB 1168 PETRKPINECKESEKDI-----VAKDGVVGEQSEIVSQPENSXDL-- 1209
QY 818 DFSESVSGTKFEEDHLSHYSPWSCGTIGSCINAIDSEPKDVIANGNAVIMLDLSDGVKRR 877
DB 1210 ----STSANKETSIVEQVAE-----KSLIAEKPDVRRSSS---IDL-NGD---- 1247
QY 878 VHLFETQRTKEEDPIIPFSDGPIISKWGAISRSRTGHTTDPVQATASQSATKPISV 937
DB 1248 --LIEAVATVLDD---SDDESPIPS-----TSVSSSVLDSQTQSEAQQDAINRVLGF 1295
QY 938 SDYVYVNAVDSRWSSYGNATSSAHYVERDFIVTDLSGHRKHSSTGDLSSLLELQOAKS 997
DB 1296 SKQCBETNF--SIWTGFENECEPEVIDI--NDNVLAAILPDE-----EPTTEE 1341
QY 998 NSLLQREANALAMQKWNLSDEGRHLTLNLLSKELFANGELQSDYTEDATDTKPDRI 1057
DB 1342 NDVYTE-----IEDKARDFGCFQKPTLLTICQVEAYFSEHHDYVGSVVD-----DM 1389
QY 1058 ELELSALOT---DEPDGQSEPIEELIDQLGSISSNDQLNG-----MAV 1099
DB 1390 ELRKIFKETYGGISIPEDHKIIEEVCSFQFVGHVFDALIGSKLEFTYTIMSNIDRKIEK 1449
QY 1100 ENGHVPQHQKPPKQKQSLGEDHVLLEEQTILP 1135
DB 1450 ENDRLYQHYSSRPADV-----GVENFYERFEKEILP 1480
RESULT 12
F86387
probable Pto kinase interactor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F86387
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F86387
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-760 <STO>
A:Cross-references: UNIPROT:Q9C660; GB:AE005172; NID:gl1079512; PIDN:AG29223.1; GSPDB:
C:Genetics:
A:Map position: 1
Query Match 3.1%; Score 190.5; DB 2; Length 760;
Best Local Similarity 25.4%; Pred. No. 0.0089;
Matches 85; Conservative 29; Mismatches 116; Indels 105; Gaps 15;

QY 531 GONAGPSADSVTENKIGSPKPTVSNVAATAGSNVGTSLNSVPOKSSPFL----- 583
Db 35 GONATSPTRFNP-----GNPPT--TNTPAQSSPPE--TPUSSPPPEPPSPSLTGP 86
QY 584 -TRVPYPPHSENIQFQDPTQIPFEV-----QYPTQGYTPPPPTVPAGVAPCV 634
Db 87 PTTPVSPPEPS-----PPPLTEAPPANPVSSPPPESSPPPPPTTEAPPTTPTS 140
QY 635 RVRGNVPESELPLPASMADHYSTFSRDRMNSPVQPP---PPQYGPVPPVPSGMY 691
Db 141 PSPPTNPPPPPSPP--SLPAPDPPSNPLPPPKLVPPSHSPRHLPSPPASEIPPPP----- 195
QY 692 APVDSRRWRPMPVQRDDIIRSNSLPPMDVMHSSVYQTSLERY----- 736
Db 196 -----RHLPSPASERP-----STPPSDSEHPGPPPPGHKREQPPPGSKRPTSP 243
QY 737 -NSLDGYVAVACQPPSEPTTVPLEPREPCGHLKTSCEEQIRRKPDQWQAQYHTQKAPLVSS 795
Db 244 PPSDSKRPVHSPSPSPPEITLPPKP-----SPD-----PLPSN 278
QY 796 TLPVATQSP-----TPPSPLFSVDFRADSESVSG 825
Db 279 SSSPPTLLPSSVWSPSPSPP-----RKSVSG 304
RESULT 13
S60402
protein kinase CLA4 (EC 2.7.1.-) - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein N0450; protein YNL0450; protein YNL298w
C:Species: Saccharomyces cerevisiae
C:Dates: 27-Apr-1996 #sequence revision 17-May-1996 #text_change 16-Aug-2004
C:Accession: S60402; S63274; S53103
R:Maurer, K.C.T.; Urbanus, J.H.M.; Planta, R.J.
Yeast 11, 1303-1310, 1995
A:Title: Sequence analysis of a 30 kb DNA segment from yeast chromosome XIV carrying a X
C. and a novel putative serine/threonine protein kinase gene.
A:Reference number: S60394; MUID:96132033; PMID:8553702
A:Accession: S60402
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-842 <MAU>
A:Cross-references: UNIPROT:P48562; EMBL:U231084; NID:g1050853; PIDN:AAC49100.1; PID:g105
R:Maurer, C.T.C.; Urbanus, J.H.M.; Planta, R.J.
submitted to the EMBL Data Library, November 1994
A:Reference number: S63266
A:Accession: S63274
A:Molecule type: DNA
A:Residues: 1-842 <MAU>
A:Cross-references: EMBL:Z71574; NID:g1302389; PIDN:CAA96216.1; PID:e239731; PID:g130239
R:Chrckova, F.; Nasmyth, K.
submitted to the EMBL Data Library, November 1994
A:Description: STE20-like protein kinases are required for cytokinesis.
A:Reference number: S53103
A:Accession: S53103
A:Molecule type: DNA
A:Residues: 1-389,'I',391-842 <CVR>
A:Cross-references: EMBL:X82499; NID:g732943; PIDN:CAAS7879.1; PID:g732944
A:Experimental source: strain K1107
C:Genetics:
A:Gene: SGD:CLA4; ERC10
A:Cross-references: SGD:S0005242; MIPS:YNL298w
A:Map position: 14L
C:Superfamily: pleckstrin repeat homology; protein kinase homology
C:Keywords: ATP; phosphotransferase
F:544-825/Domain: protein kinase homology <KIN>
Query Match 3.1%; Score 190.5; DB 2; Length 842;
Best Local Similarity 19.9%; Pred. No. 0.01;
Matches 153; Conservative 104; Mismatches 275; Indels 237; Gaps 37;
QY 272 SYEAL---RREHDAQIVHIAWEAGLRISPEQWSSLLYGLAHKSHMQSIIDKLQSPESA 328

Db 90 SYVALYKNDKQNDAILKIPLTISIISVSRTOQLKOYCF--ELVRCSDRNSV----- 137
QY 329 KSVQELTIVLORTGDPANLRLRPHLELLANIDPNDAVSTWEOLENAVAVKTV--VH 386
Db 138 -----SSGSSSLN-----VSDSNSKSIYIATKTESDLH 168
QY 387 GLVDFI-----QVSRKGHETPOPNMSKYKTSMCRDLRQOGGCPGCTNCTF 433
Db 169 SWLDAIFAKCPLLSLGVSSPTFTHKVHVGFDPETGSFV-----GNPTWEKLL 216
QY 434 AHSQBELEKYRLRNKKINATVTPTELLNKVGNVNTVTTTAGNVISVIGSTETTGKIVPST 493
Db 217 KHS-----RITGDMN-----NNSAA-----VIQVL-----QFYQBY 243
QY 494 NGISNAENSVQLISRTDSTLRALETV---KKVGKVGANGONAAGPSADSVTENKIGSP 550
Db 244 NGAGNPTNTLDKPOSGETSSQSLPNSYNDNKLNNNSVNSKSSGVSSMWVSQRKTSP 303
QY 551 PKT-----PVSNVAATSAGSNVGTSLNSVPKQSPFL-----TRVP 587
Db 304 PNTKSPVSLGSGSLPPIKLTPTS--QSNIPHLQNVFNQYPKVRNGHSPINGQFPKGP 361
QY 588 VYPHSE-NIQVFQDPTQIPFEVQYP--QTGYYPPTTVPAGVAPCVPREVRNNVPE 644
Db 362 MHPNNSQRSLQOQQOQQOQKQHQYFVHQGPSFSPSPSPLNYPHMIN--PY 419
QY 645 SSLPASMPYADHYSTFSRDRMNSP---YQPPPPQYGPVPPVPSGMYAPVVDSSRI 700
Db 420 SKQPOSPLSSOSTONQAIAPYQANSPTAAHFQ---PORTAPKPPIS---AP----- 465
QY 701 WRPPMYQRDDIIRSNLPPMDVMHSSVYQTSLERYNSLDGYYSVACOPPSPTTVPPLP 760
Db 466 -RAPYPSNQNSATNTHQVPAPKNDQSTQNR-----QAPKPPDADVAQP 510
QY 761 REPCGHLKTSCEEQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPPTPPSPFLFSVDFRADFS 820
Db 511 ---GGV-----AKPKPARPTMTSTABTMSKLKVTVNA--DPSQCFKVEKA--G 553
QY 821 EHSVSGTKF--EDHLSHYSPWSCGTCGICINAIIDSEP-----KDVTA 860
Db 554 QGASGVVLAERTHI-----PTESNMIELINNDID--EPHVGDKVAIKQWVLSQKRELIV 608
QY 861 NSNAVLMOLDSDGVKRRVHLPTQRTKEE-DPIIPFSDGPIISKWGAISRSRGTGYHTT 919
Db 609 NEILVWKO---SRHKVNVFLYLAERTDDDLWVMEFMEG-----GSL-----T 649
QY 920 DPVQATASQGSATKPIVSVDYVYVYVNAVDSRWSSYGNATSAHYVERD 968
Db 650 DIIEENPTNDNSHSLT-EPQIAYI---VRETQGLKFLHDKHIIHRD 693
RESULT 14
T32008
hypothetical protein K10G6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32008
R:Davidson, S.; Wohlmann, P.; Mullen, G.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid K10G6.
A:Reference number: Z21111
A:Accession: T32008
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1819 <DAV>
A:Cross-references: UNIPROT:O16625; EMBL:AF016669; PIDN:AAB66098.1; GSPDB:GN00020; CESP
A:Experimental source: strain Bristol N2; clone K10G6
C:Genetics:
A:Gene: CESP:K10G6.3
A:Map position: 2
A:Introns: 41/2; 191/2; 556/2; 701/2; 1645/2; 1700/2; 1768/1; 1797/3

This Page Blank (uspto)

APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: Encoding the Same
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 48
LENGTH: 1939
TYPE: PRT
ORGANISM: Homo sapiens
NAME/KEY: VARIANT
LOCATION: (1)...(1939)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-48

Query Match 3.1%; Score 193.5; DB 4; Length 1939;
Best Local Similarity 19.2%; Pred. No. 7.3e-06;
Matches 256; Conservative 168; Mismatches 527; Indels 379; Gaps 57;

QY 114 GCGVASLQNSALSRPMOKLVLYNLCVBEGRVR-AMRAARSLGERTVTTELILQHN 172
DB 660 GQFSSSLTGVSSSQBIHQHPQOQGIQTAPQQTVOYLSQTSSEATTAQVSPQQA 719
QY 173 POOLSANLWAAVRARCCQFLGPA-----MOEALKLVLLALEDDGSAKRLVLFVVRQL 227
DB 720 PQVLPOVAGKSTQGVSAEPAEPAVAQPOATQPTTLASSVDSAHSD---VASGMSDG 776
QY 228 EPRFPOAKSTISHVQLLYRASCVKVTRKDEBSSLMOLK-----EPRFSVEALRREHD 281
DB 777 NENVPSSSGRHEGRVTKHYRKS-VRSRSRHEKTSRPKRLILNVSNKGRVVECOLETHN 835
QY 282 AQVHTAMEAGLRISPEQWSSLLY-GDLAKHSHMQSIIDKLOSPEFSAKSVQELITVLQR 340
DB 836 RKWTFKFDLD-GDNPEEITATIMVNDFFILATERESFVDQVREITEKADEMLSEDSVSEP 894
QY 341 TGDPAANLRLRPHLELLANID-----PNPDVAVSPTWEQLENAMVAVKTV 384
DB 895 EGDG-----LESLOGKDDYGFSGSKLGEFKQPIPASSMP--QIGIPTSLTQV 944
QY 385 VHGLVDFI QYGRKHGHTPOQPN-----SKYKTSKCRDLRQGGCGPRGTNCTFAHS 436
DB 945 VH-----SAGRRFTVSPVESRLRESKVPFSEITDTVAASATAQSPG-----MNLSHS 991
QY 437 QBELEKVR-----LRNKKI-----NATVTRTP-----LLNKVGNVNTVTTA--- 473
DB 992 ASLSLQQAFAELRRAQMTGEPNTAPNFSHTGTPFPVVPVPLSSITAGVPTTAAATAPVP 1051
QY 474 -----GNVISVIGSTETTKGIVPSTNGISNAENS-----VSQLS 508
DB 1052 ATSSPPNDISTSVIQEVT---VPTTEGIAGVATSTGVVTSVSGGLPIPPVSESPVLSVVS 1108
QY 509 RSTDSTLRALETVKVKGKVGAN-----QONAAGPSAD-SVTENKIG-----SPPKTPVSN 557
DB 1109 SITIPAVSISTTSPSLQVPTSTSEIWSSTALYPSVTVSATSAAGGSTATPGKPPAV 1168
QY 558 VAATSAGPNVGTLSNV-----PQKSSPLTRVVPVPPHSENIQVF 599
DB 1169 VSQQAAGSTVGTATLSVSTTTSFPSTASQLSITLSSTSTPTLAETVVVVAHSLD-KTS 1227
QY 600 QDPRTQIPREVQYQOTGYPPPP-----TVPAGVAPCV-PRFVRNNVPESLPPASMP 653
DB 1228 HSTTGLAFLS-SAPSSSSPGAGVSSYISQPGGLHPLVIPSVIATPTLPQAGTSTP 1286
QY 654 YADHYSTFSPDRMNSSPQPPPPQYGPVPPVPSGM-----YAPVYDS 697
DB 1287 L-----LPQVPSIPLVQPVANVPAVQOTLIHSQPOPALLPNQHPTHCPEVDS 1334

QY 698 RRIWRPMYQRDDI-----IRS--NSLPPMDVMHSSV-YOTSLRERYNSLDGYYSVA-- 746
DB 1335 DT--QPKAGIDDIKTLEKRLRSLFSEHSSSSGAQHASVLETSLVIESVTVPITPATA 1392
QY 747 -----CQPPSE-PRITVPLPREPC---GHLKTSCEQIIR-KPDQWAQVHTQQA 790
DB 1393 PSKLLTSTTCLPTNLPLGTVALPVTVPVTPGQVSTPVSTTSGVKPGT-----APSKP 1448
QY 791 PLVSS-TLPVATQSPPTPSPPLFSVDPRADFSSVSGTKFEEDHLSHYSPWSCGTIGSCIN 849
DB 1449 PLTKAPVLPVGTGELP-----AGTLPSS--E 1470
QY 850 AIDSEPKDVIANSNAVLMDLDS-----GDVKKRRVHLFETQRTK----- 888
DB 1471 QLPPFPGLTQSQOQPLEDLDAQLRRTLSPEXITVTSVAVGPVMAAPTAITEAGTQPKG 1530
QY 889 -----BEDPIIPSPDGPILSKWG-----AISRSSTGHTTDPVQATASQGSATKPIVS 938
DB 1531 VSQVKEGVLATSSGAGVFKMGRFQVSVAAQGAQKGNKSDAKSVHFESSTSESSVLS 1590
QY 939 DIVPYVNAV-----DSRWSSYGNKETS-SAHYVERDRFIVTDLS---G 977
DB 1591 SSSPESTLVKPEPNGITIPGISSDVPESAHTTASEAKSDTCQPTKVGRFQVTTANKVG 1650
QY 978 HRKHSSTGDLLS-----LELQQAQNSLLQLQREANALAMQOKWN----- 1016
DB 1651 RPSVKTEKIDTKKEGVPASPPFMDLEQAVLPAVPIKKEPELSEPSHLNGPSSDPEA 1710
QY 1017 -----SLDEGR---HLTNLNSKIELRN-----GELQSDY-TEDATTKPDRDIELEL 1061
DB 1711 AFLSRDVEDGSGSPHQLSKSLPSQNSLSNSFSSYMSSDNESDIEDIDLKLEL 1770
QY 1062 SALDDEPDGQSEPIEELIDLGISSSQNDQLN-----GMAVENGHPVQHQKEPK 1114
DB 1771 RRL-----RDKHLKEIQLQSRQKHEIESLYTKLGKVPVPAVIPPAAAPLSGRRRRPTK 1823
QY 1115 QK-----KQSLGEDH-----VILEEQKTLIPVTSCFSPQLPVSISNASCLP 1155
DB 1824 SKGSKSSRSSSLGNKSPQLSGNLSQSAASVLHPQOTLHPGPN-----IPESQGNOLLOP 1878
QY 1156 ITTSVSAGNL 1165
DB 1879 LKPSPSDNL 1888

RESULT 9

US-09-854-856-16
Sequence 16, Application US/09854856
Patent No. 6541252
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Hilbun, Erin
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
FILE REFERENCE: Encoding the Same
CURRENT APPLICATION NUMBER: US/09/854,856
CURRENT FILING DATE: 2001-05-14
PRIORITY FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 64
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 1999
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1999)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-16


```

341 TGDPAANLRLRPHLELLANID-----PNPDVSPPTWQLENAMVAVKTV 384
895 EGDQG-----LESLOKDDYGFSGSKLEGEFKQIPASSMP--QQIGIPTSSLTQV 944
385 VHGLVDFFIONYSKRGHETPOQPN-----SKYKTSMDRLRQGGCPRGTCNCTFAHS 436
945 VH-----SAGRRFIVSPVSRLESKVFSEITDVAASTAQSPG-----MNLSHS 991
437 QBELEKYR-----LRNKKI-----NATVTFPP-----LLNKGVNNTVTFTA--- 473
992 ASLSLQQAFAFSELRRAQMTGENTAPPNFSHTGPTFPVVPFLSSLAGVPTTAATAPVP 1051
474 -----GNVISVIGSTETTKIVPSTNGISNAENS-----VSQLS 508
1052 ATSSPNDISTSVIOSEVT---VPTTEGIAGVATSTGVVTSGLPIPPVSESVLSVVS 1108
509 RSTDSTLRALETIVKVKGVKYGAN-----QONAGPSAD--SVTENKIG---SPPKTPVSN 557
1109 SITIPAVSISTTSPSLQVPTSTSEIVSVSTALYPSVTVSATSAAGSGSTATPGPKPPAV 1168
558 VAATSAGPSNVGTENSV-----POKSSPFLTRVVPVPPHSENIQYF 599
1169 VSQAAGSTTVGATLTSVSTTSPSTASQLSIQLSSSTSTPTLAETVVVSAHSLD-KTS 1227
600 QDPRTQIPPEVPOYPOTGYTPPPP-----TVPAGVAPCV-PRFVRNNVPESLPPASMP 653
1228 HSTTGLAFLS-SAPSSSSPGAGVSSYISQPGHLPLVPSVIASTPILPOAGPTSTP 1286
654 YADHYSTFPRRMNSPYQPPPPQYGPVPPVPSGM-----YAPVDS 697
1287 L-----LPQVSPITPLVQVANVPAVQVTLIHSQPQALLPNQPHTHCPEVDS 1334
698 RIWRPMPYQRDI-----IRS--NSLPMDVMHSSV-YOTSLRERYNSLDGYSAV-- 746
1335 DT--QPKAPCIDDIKTLEKRLSFLSEHSSGQAHSLETSVLVISTVTPGIPITAVA 1392
747 -----COPPE-PRTTVPLPREPC---GHLKTSCEBQIR-CPDQMAQVHTQKA 790
1393 PSKLLTSTSTCLPPTNPLGTVLVPVTPVPGVSTVPSITTSVKPCT-----APSKP 1448
791 PLVSS--TLPVATQSTPPSPPLFSVDFRADFSESVSGTKFEEDHLSHYSPWSCGTIGSCIN 849
1449 PLTKAPVLPVGTLP-----AGTLPS--E 1470
850 ADISEPKDVIANSNVIMLDS-----GDYKRVHLFETQRTK---- 888
1471 QLPPFPFGLSQTQSQPLEDLDAQLRRLTGPXITVTSVAGVPVMAAPTATEAGTQPKG 1530
889 ----BEDPIIPFSDGPIISKWG-----AISRSRTGYHTTDPVQATASQGSATKPIVS 938
1531 VSQVKEGVLATSSGAGVFKGFRFQVSVADGAQKQKSEDAKSVHPESSTSESIVLS 1590
939 DYVPYVNAV-----DSRWSSYGNBATS-SAHYVBRDRPIVTDLS---G 977
1591 SSSPESTLVKPEPENGITIPGISSDVPESAHTTASEAKSDTGQPTKVGRFOVTTANKVG 1650
978 HRKHSSTGDLSS-----LELOAKNSLLLOREANALAMQKWN----- 1016
1651 RFSVSKTEDKITDITKKEGVPASPPFMDLEQAVLPAVIPKEKPELSEPHLNGPSSDPEA 1710
1017 -----SLDEGR---HLTNLLSKIEILRN-----GELQSDY--TEDATDKPRDLELEL 1061
1711 AFLSRDVGSGFPHGPHQLSSKLSFQNSLSQSLNSFNSSYSSNESDIEDIEDLKLLEL 1770
1062 SALDTEBPQCGSEPIBEILDILQGISQNDQLLN-----GNMAVENGHIPVQOQKEPPK 1114
1771 RRL-----RDXHLKEIQDLQSRQKHEIESLYTLKGVKVPVAVIIPPAPLSGRRRRPTK 1823
1115 QK-----KQSLGEDH-----VILBEQKTLVPVTCFQOPLPVISINASCPL 1155
1824 SKGSKSRSSSLGNKSPQLSGNLSGQSAASVLHPQOTLHPGPN-----IPESQNLQLOP 1878
1156 ITTSVSAGNL 1165

```

```

Db      1879 LKPSPSSDNL 1888
RESULT 11
US-09-854-856-14
; Sequence 14, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 2136
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2136)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-14
Query Match      3.1%; Score 193.5; DB 4; Length 2136;
Best Local Similarity 19.2%; Pred. No. 8.6e-06;
Matches 256; Conservative 168; Mismatches 527; Indels 379; Gaps 57;
QY      114 GKGVASLNSQALSALRPMQKLVLYNVCQLVEEGRVR-AMRAARSIGERTVTTELLIQQHN 172
Db      720 GQPSSSSLTGVSSSQPIQHPOQQOQIQQTAPPTQVQVSLQTSSTSEATTAQPVSPQA 779
QY      173 POOLSANLWAVARACQFLGPA-----MOEALKLVLLEDEGSAKSLKVLVLFVQRL 227
Db      780 PQVLQVQVAGKOSTQGVQVAPAEVAVAPQATQPTTLASSVDSAHSD---VASGMSDG 836
QY      228 EPRFPAQKSTSGHVVOLLVYRASCFTKTRDSDSLMOLK-----EERFSYEALRREHD 281
Db      837 NENVPSSSGRHGRTTKHYRKS-VRSRSHKTSRPLRLILNVNSKGRVVEQCLETHN 895
QY      282 AQIVHIANEAGLRISPEQWSSLLY-GDLAKHSHMOSIIDKLOSPESFAKSVQELTIVLQR 340
Db      896 RKMVTFKFDLD-GDNPEEATIMVNNDFILATERSFVDQVREIIEKADEMLSEDSVSEVP 954
QY      341 TGDPAANLRLRPHLELLANID-----PNPDVSPPTWQLENAMVAVKTV 384
Db      955 EGDQG-----LESLOKDDYGFSGSKLEGEFKQIPASSMP--QQIGIPTSSLTQV 1004
QY      385 VHGLVDFFIONYSKRGHETPOQPN-----SKYKTSMDRLRQGGCPRGTCNCTFAHS 436
Db      1005 VH-----SAGRRFIVSPVSRLESKVFSEITDVAASTAQSPG-----MNLSHS 1051
QY      437 QBELEKYR-----LRNKKI-----NATVTFPP-----LLNKGVNNTVTFTA--- 473
Db      1052 ASLSLQQAFAFSELRRAQMTGENTAPPNFSHTGPTFPVVPFLSSLAGVPTTAATAPVP 1111
QY      474 -----GNVISVIGSTETTKIVPSTNGISNAENS-----VSQLS 508
Db      1112 ATSSPNDISTSVIOSEVT---VPTTEGIAGVATSTGVVTSGLPIPPVSESVLSVVS 1168
QY      509 RSTDSTLRALETIVKVKGVKYGAN-----QONAGPSAD--SVTENKIG---SPPKTPVSN 557
Db      1169 SITIPAVSISTTSPSLQVPTSTSEIVSVSTALYPSVTVSATSAAGSGSTATPGPKPPAV 1228
QY      558 VAATSAGPSNVGTENSV-----POKSSPFLTRVVPVPPHSENIQYF 599

```


Db 1436 -----AGTLPS---EQLPPFPGLSTOSQOQPLEDLDAQLRRLTSPXITVTSAVGPVSMAP 1489
Qy 880 LPEQRTK-----EDPIIPFSDGPIISKWG-----AISRSRTGYHTDPVQAT 925
Db 1490 TAITAGTQPKGVQKGVKPVLTSSGAGVFMGRFQVSVAAAGKQKSEDAKSV 1549
Qy 926 ASQGSATKPIVSVDYVYVNAV-----DSRWSSYGNBATS-SAHYVER 967
Db 1550 HFESSTSESSVLSSSPSTLVKPEPNGITIPGISDVPESAHTTASEAKSDTQPTKV 1609
Qy 968 DRPIVTDLS---CHRKHSSTGDLSS-----LELOAKSNLSLLQREANALAM 1011
Db 1610 GRFOVTTTANKVGRFSVSKTDKITDTKKEGVPASPPFMDLBOAVLPAVIPKKEPESL 1669
Qy 1012 QQKWN-----SLDEGR---HLTLNLSKEIELRN-----GELOSDY-TEDA 1048
Db 1670 PSHLNGPSSDPAEAFLSRDVDDGSGSPHQLSKSLPQNLQSLSNFSNYSN 1729
Qy 1049 TDTKPDRIEELSALDTPDQGPSEPIEILDIQLGISNDQLLN-----GMAVEN 1101
Db 1730 ESDIEDDLKLELRL-----RDKHKEIQDLSRQKHEIESLYTLKGVPPAVIIPP 1782
Qy 1102 GHVVOHQKEPKQK-----KQSLGEDH-----VILEEQKTLPTVSCFSQ 1142
Db 1783 AAPLSGRRRPTKSGSKSRSSSLGNKSPQLSGNLGQSAASVLPQOQLHPGPN---- 1838
Qy 1143 PLPVSISSNACLPITTSVAGNL 1165
Db 1839 -IPESQNLQLOLPKPSDDNL 1860

RESULT 13

US-09-854-856-32
; Sequence 32, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 1971
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)---(1971)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-32

Query Match 3.1%; Score 193; DB 4; Length 1971;
Best Local Similarity 19.3%; Pred. No. 8.2e-06;
Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;
Qy 160 ERTVTEILIQHNPQQLSANLWNAVRARGCQFLGPA-----MQEBAKLVLALDQSGAL 214
Db 739 EATTAQVSPQAPVLPQVSAQKQSTQGVSAQVAPAEVAVAFQATQPTTLASSVDSAH 798
Qy 215 SRKVLVFLVQRLPEPQASKTSIGHVQVLLYRASCFTKEDSSLMQLK-----E 268
Db 799 SD---VASGMSDGNENPVSSGRHEGRTTKRHYRKS-VRSRRHEKTSRPLKILNVSNK 854
Qy 269 EPRSVALREHDAQIVHIAEAGLRISPROWSLLY-GDLAKSHMQSIIDKLQSPESF 327

Db 855 GDRVVECOLETHNRKMWTFKFDLD-GDNPEIATIMVNDPFLAIATERBSFVDQVREIEIK 913
Qy 328 AKSVOELTIVLQRTGDPANLRLRPHLELLANID-----NPDVAVSPTM 371
Db 914 ADEMLSDVSVEPEGDQ-----LESLOQDDYGFSGSKLEGEFKQPIPASSMP-- 963
Qy 372 EOLENAMAVKTVHGLVDFIQNYSRKHEHETPOPOP-----SKYKTSMDRLRQOG 423
Db 964 QOIGIPTSLTQVWH-----SAGRRFIVSPVPSRLRESKVFSEITDVTAAATASP 1016
Qy 424 GCPRTGNTCTFAHSQSELEKYR---LRNKKI-----NATVRTPP-----LLNKV 463
Db 1017 G-----MNLHSASSLSLQOAFSELRAQMEGTAPPNFSHTGTFPVVPPFLSSIA 1070
Qy 464 GYNNTVTITTA-----GNVISVIGSTETTKIVPSTNGISNAENS----- 502
Db 1071 GVPTTAATAAPVATSPSPNDISTSVIQSEVT---VPEEGIAGVATSGVVTSGGLPIP 1127
Qy 503 -----VSQILSRSTDSTLRALETVKVKGVKGAN-----GQNAAGPSAD-SVTENKIG 548
Db 1128 PVSESPVLSSVVSSITIPAVVSIISTTSPSLQVPTSTSEIVVSSTALYPSVTVSATASAG 1187
Qy 549 ---SPPKTPVSNVAATAGPSNVCTELNSV-----POKSSPFLTRV 586
Db 1188 GSTATPGKPPAVVQQAGSTTUGATLTSVSTTSPSTASQLSIQLSSSTSTPTLAET 1247
Qy 587 PVYPPHSENIQYFQDPTQIPFEVPOYPTQGYYPPPP-----TVPAGVAPCV-PRFVRN 640
Db 1248 VVVSASHLD-KTSHSSTTGLAFSL-SAPSSSSSPGAGVSSYISQGLHPLVIPSVIAT 1305
Qy 641 NVPESSLPASMPYADHYSTSPRDRMNSSPQPPPPQPYGVPVPPVPSGM----- 690
Db 1306 PILQAAGTSTPL-----LPQVPSIPPLVQFVANVPVAVQOQLIHSQOPALL 1353
Qy 691 -----YAPVYDSRIWPPMWYQDDI-----IRS--NSLPPMDVMHSSV-YOTSLRER 735
Db 1354 PNQPHTHCEVDSOT--QPKAPGIDDIKTLBEKLSLSEHSSSSGAQHASVLESLVIE 1411
Qy 736 YNSLDGYYSVA-----CQPPSE-PRITVPLPREPC---GHLKTSCEEQIRR- 777
Db 1412 SITVTGITTAVAPSKLLTSTTCLPTNLPLGTVALPVTVPVTPGVQVSTPSTTSGV 1471
Qy 778 KPDQWAQYHTQKAPLVSS-TLPVATQSPPTPSPPLFSVDPRADPSESVCSTKFEEDHLSHY 836
Db 1472 KPCT---APSKPPLTKAPVLPVGTLP----- 1495
Qy 837 SPWSCGTIGSCINADSEPKDVIANNAVLMDLDS-----GDKRRVH 879
Db 1496 ----AGTLPS---EQLPPFPGLSTOSQOQPLEDLDAQLRRLTSPXITVTSAVGPVSMAP 1549
Qy 880 LPEQRTK-----EDPIIPFSDGPIISKWG-----AISRSRTGYHTDPVQAT 925
Db 1550 TAITAGTQPKGVQKGVKPVLTSSGAGVFMGRFQVSVAAAGKQKSEDAKSV 1609
Qy 926 ASQGSATKPIVSVDYVYVNAV-----DSRWSSYGNBATS-SAHYVER 967
Db 1610 HFESSTSESSVLSSSPSTLVKPEPNGITIPGISDVPESAHTTASEAKSDTQPTKV 1669
Qy 968 DRPIVTDLS---CHRKHSSTGDLSS-----LELOAKSNLSLLQREANALAM 1011
Db 1670 GRFOVTTTANKVGRFSVSKTDKITDTKKEGVPASPPFMDLBOAVLPAVIPKKEPESL 1729
Qy 1012 QQKWN-----SLDEGR---HLTLNLSKEIELRN-----GELOSDY-TEDA 1048
Db 1730 PSHLNGPSSDPAEAFLSRDVDDGSGSPHQLSKSLPQNLQSLSNFSNYSN 1789
Qy 1049 TDTKPDRIEELSALDTPDQGPSEPIEILDIQLGISNDQLLN-----GMAVEN 1101
Db 1790 ESDIEDDLKLELRL-----RDKHKEIQDLSRQKHEIESLYTLKGVPPAVIIPP 1842
Qy 1102 GHVVOHQKEPKQK-----KQSLGEDH-----VILEEQKTLPTVSCFSQ 1142
Db 1843 AAPLSGRRRPTKSGSKSRSSSLGNKSPQLSGNLGQSAASVLPQOQLHPGPN---- 1898

```
QY 1143 PLPVSIASCLPITTSVSAGNL 1165
Db 1899 -IPESGQNQLLOPLKPSSSDNL 1920

RESULT 14
US-09-854-856-62
; Sequence 62, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 2048
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2048)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-62

Query Match 3.1%; Score 193; DB 4; Length 2048;
Best Local Similarity 19.3%; Pred. No. 8.8e-06;
Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

QY 160 ERTVELILQHNPQOLSANLWAVRARCQFLGA-----MQEALKVLVLALEDGSL 214
Db 679 EATTAQVSPQAPQVLPQVSAQKOSTQGVSAFAEPVAVAPQATQPTTLASSVDSAH 738
QY 215 SRKVLVLFVQRLPFPQAKTSIGHVVLRYRASCVKVTKRDBDSSLMOLK-----E 268
Db 739 SD---VASGMSGNGENVSPSSRGHRTTKRYRKS-VRSRSHKTSRPKRLILNVNK 794
QY 269 EFRSVEALRREDAQIVHTIAMBAGLRISPEQWSLLY-GDLAKHSHMOSIIDKLOSPSP 327
Db 795 GDRVVECOLETHNRKMTVFKFDLD-GDNPEEATIMVNDFFILATERESFVDQVREIEK 853
QY 328 AKSVQELTIVLQRTGDPANLRLRPHLELLANID-----PNPDVSPTW 371
Db 854 ADEMLSEDVSPPEGDQ-----LESLOGKDDYGFSGSKLEGEFKQPTIPASSMP-- 903
QY 372 EOLENANWAKTVTHGLVDFIYNGRKHGHTPOPN-----SKYKTSRCRDLROQG 423
Db 904 QQIGPITSLTQVH-----SAGRRFIVSPVPSRLRESKVFSEITDVAASTAQP 956
QY 424 GCPRGNTCTFAHSEBLEKYR-----LRNKI-----NATVRTFP-----LLNKV 463
Db 957 G-----MNLSSHASLSLQAFSELRRRAQTEGPNATPPNFSHTGPTFPVVPPLSSIA 1010
QY 464 GUNNVITTTA-----GNVISIGSTETTKIVPSTNGISNAENS-----502
Db 1011 GVPTTAAATAPVATSPSPNDISTSVIQSEVT---VPTBEGIAGVATSTGVVTSGLPLP 1067
QY 503 -----VSQLSRSRSTSLRALETVKVKGVKAN-----GQNAAGPSAD-SVTENKIG 548
Db 1068 PVSEPVLSVSSVITIPAVVISITSPSLQVPTSTSEIWSSTALVPVTSATASAG 1127
QY 549 -----SPKPTFVSNVAATAGSPNVGTELNSV-----PKKSSPPLTRV 586
Db 1128 GSTATPGPKPPAVVSOQAAGSTTVGATLTLSVTTTFFPSTASQLSITLSSTSTPTLAET 1187

QY 587 PVYPPHSENIQYFQDPTQIPEVEPQYQYQYPPPP-----TVPAGVAPCV-PRFVRSN 640
Db 1188 VVSAHSLD-KTSHSSTTGLAFSL-SAPSSSSPCAGVSSYISQPGGLHPLVIPSVIAS 1245
QY 641 NVPESSLPASMPYADHYSTFSPRDRMNSPYQPPPPQYQYQYPPVPPVPSGM----- 690
Db 1246 PILQOAGPTSTPL-----LPQVFSIPLVQVAVANPAVQOTLIHSQOQPALL 1293
QY 691 -----YAPVYDSRRIWRPPMYQRDDI-----TRS--NSLPPMDVMISSV-YQTSLE 735
Db 1294 PNQPHTHCPVEDSDT--QPKAPGIDDIKTLEKRLSLFSEHSSSGAQSASVLESLVIE 1351
QY 736 YNSLDGYYSVA-----CQPPSE-PRTTVPLPREPC-----GHLKTSCEBQIRR- 777
Db 1352 STVTPIGITTAVAPSKLLTSTTCLPPTNPLGTVALPVTVPVTPGQVSTVSTTSGV 1411
QY 778 KPDQWAOYHTQKAPLVSS-TLPVATQSPTPSPPLFSVDFRADFSESVSGTKFEEDHLSH 836
Db 1412 KPGT-----APSKPPLTKAPVLPVGTLP-----GDVKKRVH 879
QY 837 SPWSCGTTGSCINAIIDSEPKDVIANNAVLMDLDS-----GDKVRRVH 879
Db 1436 ---AGTLFS--EQLPPFPFSLTQSQOQFLEDLQAQLRRLTSLPEXITVTSVAVGPVMAAP 1489
QY 880 LFEQRRTK-----EEDPIIPFSDGPIISKWG-----AISRSSRTGYHTTDPVOAT 925
Db 1490 TAITBAGTQOKGVSOVKEGPVLATSSGAGVFKGRFOVSVAAADGAQKEGKNKSEDAKSV 1549
QY 926 ASQGSATKPIVSVDYVYVNAV-----DSRWSSYGNHEATS-SAHYVER 967
Db 1550 HFESSTSESVLSSSSPESTLVKPEPGITIPGISSDVPESAHTTASEAKSDTQGTQKV 1609
QY 968 DRFIVTDLG---GHRKHSSTGDLIS-----LELOAKSNLLOREANALAM 1011
Db 1610 GRFOVTTANKVGRFSVSKTKITDTKKEGVASPPFMDLEQAVLPVAPKKEPELSE 1669
QY 1012 QOKWN-----SLDEGR--HLTNLLSKETELRN-----GELQSDY-TEDA 1048
Db 1670 PSHLNGPSSDEAPFLSRDVGSGSPHSQHSKSLPSONLSQSLNSFNSSVSSSDN 1729
QY 1049 TDTKPDRIELELSALDDEPDGQSEPIEEILDQLGISSQNDQLLN-----GMAVEN 1101
Db 1730 ESDIEDEDLKLELREL-----RDKHLKEIQDLSRQKHEIESLYTKLGKVPVAVIIPP 1782
QY 1102 GHPVOHQKPEPKQK-----KQSLGEDH-----VILEEQKTLPTVTSQFSQ 1142
Db 1783 AAPLSGRRRRTKSGKSSRSSSLGNKSPQLSGNLGQSAASVLPHPQOTLHPGPN---- 1838
QY 1143 PLPVSIASCLPITTSVSAGNL 1165
Db 1839 -IPESGQNQLLOPLKPSSSDNL 1860

RESULT 15
US-09-854-856-30
; Sequence 30, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
```

LENGTH: 2108

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)...(2108)

OTHER INFORMATION: Xaa = Any Amino Acid

US-09-854-856-30

Query Match

Best Local Similarity 3.1%; Score 193; DB 4; Length 2108;

Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

QY 160 ERTVTELIHQHNPQOLSANLWAAVRAGCQFLGPA-----NQEALKVLVLALEDGSA 214
DB 739 EATTAQPVSPQAPQVLPQVSAGKSTQGVAPAEPAVAPQATQPTTLASSVDSAH 798
QY 215 SRKVLVVFQRLPRFPQAKTSIGHVVQLLYRASCFTVTRDEDSLMQJL-----E 268
DB 799 SD---VASGMSDGNENVPSSGRHEGRTTKRHYRKS-VRSRSRHEKTSRPLRLNVSNK 854
QY 269 EFRSYEARREDAQIVHAMEAGLRISPEQWSSLLY-CDLAHSHMQSIIDKLQSPESF 327
DB 855 GDRVVECOLETHNRKMTFKFDLD-GDNPEEATIMVNNDFILATERESFVDQVREIEK 913
QY 328 AKSQBELTIVLQRTGDPANLRLRPHLELLANID-----PNPDVSPPTM 371
DB 914 ADEMLEDVSEPEGQOG-----LESLOGKDDYGFSGQKLEGEFKQIOPASSMP-- 963
QY 372 EQLENAMAVKTVVHGLVDFIQNSRKGHETPOQPN-----SKYKTSMCRDLRQOG 423
DB 964 QQIGIFTSLSITQVH-----SAGRRFIVSPVPESRLRESKVPFSEITDTVAASQAQP 1016
QY 424 GPRGNTCTFANSQEELEKVR-----LRNKKI-----NATVTRTP-----LLNKV 463
DB 1017 G-----MNLSSHASLSLQAFSELRAQMTGEPNTAPNFSHTGPTFPVVPFLSSIA 1070
QY 464 GVNNTVTTTA-----GNVISIGSTETGKIVPSTNGISNAENS----- 502
DB 1071 GVPTTAATAAPVDPATSPNDISTSVIQSEVT---VPTTEGIAGVATSGVTVSGGLPIP 1127
QY 503 -----VSQISRSRSTDLRALETVKYKVGKAN-----QONAAAGPSAD-SVTENKIG 548
DB 1128 PVSESPVLSVSVSSITIPAVVSISITSPSLQVPTSTSEIWSSTALYPSVTVSATASAG 1187
QY 549 -----SPKPTPVSNVATSGPSNVGTGLNSV-----POKSSPELTRV 586
DB 1188 GSTATGPKPPAVVSOQAAGSTTVGATLTSVSTTSPSTASQLSIQLSSSTSTPTLAET 1247
QY 587 PVYPHSENIQYFQDPRTOIPREVPOYPOTGYYPPPP-----TVPAGVAPCV-PRFVRSN 640
DB 1248 VVVSASHLD-KTSHSSTTGLAFSL-SAPSSSSFGAGVSYSIQPGGLHPLVIPSVIAST 1305
QY 641 NVPESSILPPASMPYADHYSTFSPDRMNGSPYQPPPPQPYGPVPPVPSGM----- 690
DB 1306 PILPQAAGPTSTPL-----LPQVPSIPPLVQPVANVPAVQQTLIHSQPOPALL 1353
QY 691 -----YAPVYDSRRIRWPPMYORDDI-----IRS--NSLPPMDVMHSSV-YQTSIRER 735
DB 1354 PNQPHTHCPEDVSDT--QPKAPGIDDIKTLEEKLSLFSHSSSGAQSASVLSLETSLVIE 1411
QY 736 YNSLDGYYSVA-----COPPE-PRITVPLPREPC---CHLKTSCCEEQIRR- 777
DB 1412 STVTGPIPTTAVAPSKLLTSTTCLPLGTVALPVTVPVQVSTPVSPTTSGV 1471
QY 778 KPDQWQYHTQKAPLVSS-TLPVATQSPPTPPLFSVDPRADFSVSGTKFEEDHLSHY 836
DB 1472 KPQT-----APSKPLTKAPVLPVGTLP----- 1495
QY 837 SPWSCGTIGSCINAIIDSEPKDVIANGNAVIMDLDS-----GDVKKRVH 879
DB 1496 ----AGTLPS--EQLPPPGPSLTQSQOPLDLDLAQLRRTLSPXITVTISAVGVSVMAAP 1549

Search completed: August 4, 2005, 01:49:10
Job time : 56 secs

QY 880 LFEQRTTK-----BEDPILPFGDCPIISKWG-----AISRSRTGYHTTDPVOAT 925
DB 1550 TAITAAGTQPKGVSVQKGVFVLTSSGAGVFKMGRFQVSVAAADGAQKEGKNKSEDAKSV 1609
QY 926 ASQGSATKPISVSDVYPVNAV-----DGRWSSYGNKATS-SAHYVER 967
DB 1610 HFESSTSESSVSSSSPESTLVKPEPNCITIPGISSDVPESAHTTASEAKSDTQPTKV 1669
QY 968 DRFIVTDLS---GHRKHSSTGDLJS-----LELOQAKSNLSLLQREANALAM 1011
DB 1670 GRFQVTTTANKVRGFSVSKTEDKITDTKKEGFVASPPFMDLEQAVLPAVIPKKEPELSE 1729
QY 1012 QOKWN-----SLDEGR---HLTNLLSKEIELRN-----GELOSDY-TEDA 1048
DB 1730 PSHLNGPSSDPEAAFLSRDVEDDGGSPHSPHQLSKSLPSQNLQSLSNSFNSSYMSDN 1789
QY 1049 TDTKPDRIEELSALDTPEDPGQSEPIEELIDIQLGISSQNDQLLN-----GNMVEN 1101
DB 1790 ESDIEDEDLKLELRL-----RDXHLKEIQDLSRQKHEIESLYTKLGKVPVAVIIPP 1842
QY 1102 GHPVQOQHKEPPKQK-----KQSLGEDH-----VILEEQKTLIPVTSVCSFSQ 1142
DB 1843 AAPLSGRRRRPTKSGSKSRSSSLGKNSPOLSGNLSGSAASVLPQOTLHPPGN---- 1898
QY 1143 PLPVSISNASCLPITTSVSAGNL 1165
DB 1899 -IPESGQNLQPLKPPSSDNL 1920

This Page Blank

This Page Blank (uspto)

OM protein - protein search, using sw model
Run on: August 4, 2005, 01:27:38 ; Search time 192 Seconds
(without alignments)
3176.491 Million cell updates/sec

Title: US-10-619-992-2
Perfect score: 6215
Sequence: 1 MPVQAQWTEFLSCPICYN.....MSEDKNDFLKPANGKQVNS 1191

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5483	88.2	1048	2	Q9HBD2	Q9hbd2 homo sapien
2	4412	71.0	869	2	Q9NXL1	Q9nxe1 homo sapien
3	3721	59.9	728	2	Q8N3D6	Q8n3d6 homo sapien
4	2847	45.8	540	2	Q9H5J2	Q9h5j2 homo sapien
5	2685	43.2	1136	2	Q69231	Q69231 mus musculus
6	2577.5	41.5	1109	2	Q81VE6	Q81ve6 homo sapien
7	2477	39.9	1114	2	Q6NUC6	Q6nuc6 xenopus lae
8	2293.5	36.9	506	2	Q86ST6	Q86st6 homo sapien
9	2165	34.8	419	2	Q9HBD1	Q9hbd1 homo sapien
10	1538	24.7	819	2	Q9VV48	Q9vv48 drosophila
11	1289.5	20.7	398	2	Q7PTD3	Q7ptd3 anopheles g
12	889	14.3	198	2	Q9NWN9	Q9nwn9 homo sapien
13	843	13.6	1014	2	Q45962	Q45962 caenorhabdi
14	780	12.6	177	2	Q96F27	Q96f27 homo sapien
15	611	9.8	522	2	Q8N9V1	Q8n9v1 homo sapien
16	213.5	3.4	1157	1	BBC1_YEAST	B47068 saccharomyc
17	212	3.4	5085	1	PCLO_RAT	Q9jx86 rattus norv
18	211	3.4	1918	2	Q752A6	Q752a6 ashbya goss
19	207.5	3.3	1090	2	Q9P5J0	Q9p5j0 neurospora
20	207.5	3.3	1100	2	Q7R2C6	Q7rzc6 neurospora
21	206	3.3	1289	2	Q9FLQ7	Q9flq7 arabidopsis
22	206	3.3	1611	2	Q42854	Q42854 schizosacch
23	205	3.3	1730	2	Q6FJN8	Q6fjn8 candida gla
24	204.5	3.3	9234	2	Q7KTP5	Q7ktp5 drosophila
25	204	3.3	2377	1	WNK1_MOUSE	P83741 mus musculus
26	203	3.3	786	2	Q48909	Q48909 arabidopsis
27	202	3.3	760	2	Q9T0K5	Q9t0k5 arabidopsis
28	201	3.2	1082	2	Q8VIG1	Q8vig1 mus musculus
29	200.5	3.2	728	2	Q9HFN3	Q9hfn3 candida tro
30	200	3.2	1134	2	Q632W5	Q632w5 mus musculus
31	198	3.2	684	2	Q6VQR0	Q6vqr0 homo sapien

Q6fay6 candida gla
Q9h4a3 homo sapien
Q97005 leishmania
Q65375 arabidopsis
Q78710 neurospora
Q8cf91 mus musculus
Q8cf92 mus musculus
Q18936 caenorhabdi
Q60885 homo sapien
Q9gyx7 mus musculus
Q80tw6 mus musculus
Q8la13 dictyosteli
Q43432 homo sapien
Q09085 phaseolus v

ALIGNMENTS

RESULT 1
Q9HBD2
ID Q9HBD2 PRELIMINARY; PRT; 1048 AA.
AC Q9HBD2;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Membrane-associated nucleic acid binding protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517928; PubMed=10938276; DOI=10.1074/jbc.M004461200;
RA Siess D.C., Vedder C.T., Merkens L.S., Tanaka T., Freed A.C.,
RA McCoy S.L., Heinrich M.C., Deffebach M.E., Bennett R.M.,
RA Hefenreider S.H.;
RT "A human gene coding for a membrane-associated nucleic acid-binding
RT protein.";
RL J. Biol. Chem. 275:33655-33662(2000).
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AF255303; ARG00432.1; -;
DR HSSP; P38398; 1JW7.
DR GO; GO:0009986; C:cell surface; IDA.
DR GO; GO:0005624; C:membrane fraction; IDA.
DR GO; GO:0003677; F:DNA binding; IDA.
DR InterPro; IPR000571; Znf_CCCH.
DR InterPro; IPR001841; Znf_RING.
DR Pfam; PF00642; zf-CCCH; 1.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf_C3H1; 1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
FT NON TER 1048 1048
SQ SEQUENCE 1048 AA; 116077 MW; 0487D5B5C17D20C6 CRC64;

Query Match 88.2%; Score 5483; DB 2; Length 1048;
Best Local Similarity 100.0%; Pred. No. 1.7e-265;
Matches 1048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPVQAQWTEFLSCPICYNFENVHKPSLGGSHVTCKTCLNKLHKKACPPDQTINTD 60
Db 1 MPVQAQWTEFLSCPICYNFENVHKPSLGGSHVTCKTCLNKLHKKACPPDQTINTD 60
QY 61 IDVLPVNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYKPLSGGKGVAS 120
Db 61 IDVLPVNFALLQVGAQVDPHQSIKLSNLGENKHVEVAKKVEDLALYKPLSGGKGVAS 120
QY 121 LQNSALSRPMQKRLVTLVNCQLVEEGRVAMRAARSLGERTVTTELLOHQNPOOLSANL 180
Db 121 LQNSALSRPMQKRLVTLVNCQLVEEGRVAMRAARSLGERTVTTELLOHQNPOOLSANL 180

```
QY 181 WAAVRARGCOFTGPANQOEALKLVLALLLEDGSALESKVLVLFVVRLEPRFPQASKTSTG 240
Db 181 WAAVRARGCOFTGPANQOEALKLVLALLLEDGSALESKVLVLFVVRLEPRFPQASKTSTG 240
QY 241 HVVQLLYRASCFTKVKTRDSSIMQLKEFRSYEARLRREHDAQIVHIAMEAGLRISPEQW 300
Db 241 HVVQLLYRASCFTKVKTRDSSIMQLKEFRSYEARLRREHDAQIVHIAMEAGLRISPEQW 300
QY 301 SLLYVGLLAHKSHMSIIDLQSPSPFAKSVQELTIVLQRTGDPANLRLRPHLELLANI 360
Db 301 SLLYVGLLAHKSHMSIIDLQSPSPFAKSVQELTIVLQRTGDPANLRLRPHLELLANI 360
QY 361 DNPDAVSPWQLENAWYAVKTVVHGLVDFIYONYSRKGHETPQOPNSKYKTSWCRDLR 420
Db 361 DNPDAVSPWQLENAWYAVKTVVHGLVDFIYONYSRKGHETPQOPNSKYKTSWCRDLR 420
QY 421 QGGCGPRGTNCTFAHSQOELEKYRLRNKKINATVTFPLLNKVGNNVTVTAGNVISVI 480
Db 421 QGGCGPRGTNCTFAHSQOELEKYRLRNKKINATVTFPLLNKVGNNVTVTAGNVISVI 480
QY 481 GSTETTKIVPTNGTISNAENSVSQLSIRSTDSTLRALETVKVKGVGANGQNAAGPSAD 540
Db 481 GSTETTKIVPTNGTISNAENSVSQLSIRSTDSTLRALETVKVKGVGANGQNAAGPSAD 540
QY 541 SVTENKIGSPPKTPVSNVAATAGPSNVGTLSNVPOKSSPFLTRVPVYPHSENIQYFQ 600
Db 541 SVTENKIGSPPKTPVSNVAATAGPSNVGTLSNVPOKSSPFLTRVPVYPHSENIQYFQ 600
QY 601 DPRTQIPFVPOYPOTGYPPPTVPAGVAPCVPRFVRSNNVPSSLPASMPYADHYST 660
Db 601 DPRTQIPFVPOYPOTGYPPPTVPAGVAPCVPRFVRSNNVPSSLPASMPYADHYST 660
QY 661 FSPDRMNSSPQPPPPQYGVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRSNSLPPM 720
Db 661 FSPDRMNSSPQPPPPQYGVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRSNSLPPM 720
QY 721 DVMHSSVYQTLRERNYSLDGYSVACQPPSPRTVPLPREPCGHLKTSCEEQIRRKPD 780
Db 721 DVMHSSVYQTLRERNYSLDGYSVACQPPSPRTVPLPREPCGHLKTSCEEQIRRKPD 780
QY 781 QWAQYHTQKAPLVSSTLVATQSPPTPSPLFSVDFRADFSESVSGTKFEDHLSHYSPWS 840
Db 781 QWAQYHTQKAPLVSSTLVATQSPPTPSPLFSVDFRADFSESVSGTKFEDHLSHYSPWS 840
QY 841 CQTIGSCINAIDSEPKDVIANNAVLMDLDSGDVKKRVHLFETQRTKKEEDPIIIPSDGP 900
Db 841 CQTIGSCINAIDSEPKDVIANNAVLMDLDSGDVKKRVHLFETQRTKKEEDPIIIPSDGP 900
QY 901 IISKWGAISRSRTGYHTTDPVQATASQGSATKPIISVDYVPYVNAVDSRWSSYNGEATS 960
Db 901 IISKWGAISRSRTGYHTTDPVQATASQGSATKPIISVDYVPYVNAVDSRWSSYNGEATS 960
QY 961 SAHYVERDRFIVTDLSGHRKHSSTGDLLELQQAQKSNLLOREANALAMQKNSLDE 1020
Db 961 SAHYVERDRFIVTDLSGHRKHSSTGDLLELQQAQKSNLLOREANALAMQKNSLDE 1020
QY 1021 GRHLTLNLLSKIELRNGELQSDYTEDA 1048
Db 1021 GRHLTLNLLSKIELRNGELQSDYTEDA 1048
```

RESULT 2

```
Q9NKE1 PRELIMINARY; PRT; 869 AA.
ID Q9NKE1;
AC Q9NKE1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ20301.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000308; BAA91073.1; -.
DR Genew; HGNC:8212; ORL1.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR000571; Znf.CCCH.
DR Pfam; PF00642; zf-CCCH; 1.
DR SMART; SM00356; Znf_C3H1; 1.
SQ SEQUENCE 869 AA; 96347 MW; 28105761C60B3D65 CRC64;

Query Match 71.0%; Score 4412; DB 2; Length 869;
Best Local Similarity 99.6%; Pred. No. 4.2e-212;
Matches 843; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 196 MOEALKLVLALLLEDGSALESKVLVLFVVRLEPRFPQASKTSTIGHVVQLLYRASCFTK 255
Db 1 MOEALKLVLALLLEDGSALESKVLVLFVVRLEPRFPQASKTSTIGHVVQLLYRASCFTK 60
QY 256 KDEDSSLMQLKEEPRSYEARLRREHDAQIVHIAMEAGLRISPEQWSSLLYGLAHKSHMQ 315
Db 61 KDEDSSLMQLKEEPRSYEARLRREHDAQIVHIAMEAGLRISPEQWSSLLYGLAHKSHMQ 120
QY 316 SIIDKLQSPSPFAKSVQELTIVLQRTGDPANLRLRPHLELLANIDPNDAVSPWQLE 375
Db 121 SIIDKLQSPSPFAKSVQELTIVLQRTGDPANLRLRPHLELLANIDPNDAVSPWQLE 180
QY 376 NAWAVKTVVHGLVDFIYONYSRKGHETPQOPNSKYKTSWCRDLRQOQCGPRGTNCTFAH 435
Db 181 NAWAVKTVVHGLVDFIYONYSRKGHETPQOPNSKYKTSWCRDLRQOQCGPRGTNCTFAH 240
QY 436 SOEELEKYLRNKKINATVTFPLLNKVGNNVTVTAGNVISVIGSTETTKGIVPSTNG 495
Db 241 SOEELEKYLRNKKINATVTFPLLNKVGNNVTVTAGNVISVIGSTETTKGIVPSTNG 300
QY 496 ISNAENSVSQLSIRSTDSTLRALETVKVKGVGANGQNAAGPSADSVTENKIGSPKTPV 555
Db 301 ISNAENSVSQLSIRSTDSTLRALETVKVKGVGANGQNAAGPSADSVTENKIGSPKTPV 360
QY 556 SNVAATAGPSNVGTLSNVPOKSSPFLTRVPVYPHSENIQYFODPRTQIPEFVPOYFQ 615
Db 361 SNVAATAGPSNVGTLSNVPOKSSPFLTRVPVYPHSENIQYFODPRTQIPEFVPOYFQ 420
QY 616 TGYPPPTVPAGVAPCVPRFVRSNNVPSSLPASMPYADHYSTFSPDRMNSSPYQPP 675
Db 421 TGYPPPTVPAGVAPCVPRFVRSNNVPSSLPASMPYADHYSTFSPDRMNSSPYQPP 480
QY 676 PPQYGVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRSNSLPMDVMHSSVYQTSLRER 735
Db 481 PPQYGVPPVPSGMYAPVYDSRRIRWPPMYQRDDIIRSNSLPMDVMHSSVYQTSLRER 540
QY 736 YNSLDGYSVACQPPSEPRTTVPLPREPCGHLKTSCEEQIRRKPDQAWYHTQKAPLVSS 795
Db 541 YNSLDGYSVACQPPSEPRTTVPLPREPCGHLKTSCEEQIRRKPDQAWYHTQKAPLVSS 600
QY 796 TLPVATQSPPTPSPLFSVDFRADFSESVSGTKFEDHLSHYSPWSCGTIGSCINAIDSEP 855
Db 601 TLPVATQSPPTPSPLFSVDFRADFSESVSGTKFEDHLSHYSPWSCGTIGSCINAIDSEP 660
QY 856 KOVIANSNAVLMDLSDGVKKRVHLFETQRTKKEEDPIIPPSDGLIISKWGAISRSRTG 915
Db 661 KOVIANSNAVLMDLSDGVKKRVHLFETQRTKKEEDPIIPPSDGLIISKWGAISRSRTG 720
QY 916 YHTTDPVQATASQGSATKPIISVDYVPYVNAVDSRWSSYNGEATSAHYVERDRFIVTDL 975
Db 721 YHTTDPVQATASQGSATKPIISVDYVPYVNAVDSRWSSYNGEATSAHYVERDRFIVTDL 780
QY 976 SGHRKHSSTGDLLELQQAQKSNLLOREANALAMQKNSLDEGRHITLNLSSKIEL 1035
```

Db 781 SGHRKHSCTGDLSSLSLQAKSNLLQREANALAMQKWNLSLDEGRHLTLNLLSKIEL 840
QY 1036 RINGELQ 1041
| | | | |
Db 841 RINGEVK 846
| | | | |
RESULT 3
Q8N3D6 PRELIMINARY; PRT; 728 AA.
AC Q8N3D6;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein DKFZp762N035 (Fragment)
GN Names=DKFZp762N035;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RG The German CDNA Consortium;
RA Bloecher H., Boecher M., Brandt P., Mewes H.W., Weil B., Amid C.,
RA Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RE EMBL; AL834431; CAD39091.1; -
DR GO; GO:0009986; C:cell surface; ISS.
DR GO; GO:0005624; C:membrane fraction; ISS.
DR GO; GO:0003677; F:DNA binding; ISS.
DR InterPro; IPR000571; Znf.CCCH.
DR Pfam; PF00642; zf-CCCH; 1.
DR SMART; SM00356; Znf.C3H1; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 728 AA; 80284 MW; 5B2B6C5682AA93A6 CRC64;
Query Match 59.9%; Score 3721; DB 2; Length 728;
Best Local Similarity 99.7%; Pred. No. 1.1e-177;
Matches 703; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 337 VLQRTGDPANLRPHLELLANIDPNDAVSTWEQLENAMVAVKTVHGLVDFIQNYS 396
| | | | |
Db 1 VLQRTGDPANLRPHLELLANIDPNDAVSTWEQLENAMVAVKTVHGLVDFIQNYS 60
| | | | |
QY 397 RKGHETPQOPNSKYKTSKCRDLRQGGCPRGTCNCTFAHSQBLEKYRLNKKINATVRT 456
| | | | |
Db 61 RKGHETPQOPNSKYKTSKCRDLRQGGCPRGTCNCTFAHSQBLEKYRLNKKINATVRT 120
| | | | |
QY 457 FPLLNKVGNNVTVTAGNVISVIGSTETTKIVPSTNGISNAENSVSOLISRSTOSTLR 516
| | | | |
Db 121 FPLLNKVGNNVTVTAGNVISVIGSTETTKIVPSTNGISNAENSVSOLISRSTOSTLR 180
| | | | |
QY 517 ALETVKVGKVGANGQNAAGPSADSVTENKIGSPKTPVSNVAATSAAGSNVGTSLNSVP 576
| | | | |
Db 181 ALETVKVGKVGANGQNAAGPSADSVTENKIGSPKTPVSNVAATSAAGSNVGTSLNSVP 240
| | | | |
QY 577 QKSSPELTPVVPVPHSENIQYQDPRQTQIPFEVPOYPQTGYPPPTPVAGVAPCVPRF 636
| | | | |
Db 241 QKSSPELTPVVPVPHSENIQYQDPRQTQIPFEVPOYPQTGYPPPTPVAGVAPCVPRF 300
| | | | |
QY 637 VRNNNPESLTPASMPYADHYTFSPRDMNSSPYQPPPPQYGVPPVPSPGMYAPVVD 696
| | | | |
Db 301 VRNNNPESLTPASMPYADHYTFSPRDMNSSPYQPPPPQYGVPPVPSPGMYAPVVD 360
| | | | |
QY 697 SRRIWPPMYQRDDIIRSNLPMQVHSSVYQTSRLRYSNLSLDGYYSVACQPPSPRIT 756
| | | | |
Db 361 SRRIWPPMYQRDDIIRSNLPMQVHSSVYQTSRLRYSNLSLDGYYSVACQPPSPRIT 420
| | | | |
QY 757 VPLPRPCGHLKTSCEQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPPTPPSPFVDFR 816
| | | | |
Db 421 VPLPRPCGHLKTSCEQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPPTPPSPFVDFR 480
| | | | |

QY 817 ADPSFSGTKFEEDHLSHYSPWSCGTIGSCINAIIDSEPKOVIANNAVLMDLDSGDVVR 876
| | | | |
Db 481 ADPSFSGTKFEEDHLSHYSPWSCGTIGSCINAIIDSEPKOVIANNAVLMDLDSGDVVR 540
| | | | |
QY 877 RVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLS 936
| | | | |
Db 541 RVHLFETQRTKEEDPIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLS 600
| | | | |
QY 937 VSDYVPYVNAVDSRWSSYSGNEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQQAQ 996
| | | | |
Db 601 VSDYVPYVNAVDSRWSSYSGNEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQQAQ 660
| | | | |
QY 997 SNSLLQREANALAMQKWNLSLDEGRHLTLNLLSKIELRNGELQ 1041
| | | | |
Db 661 SNSLLQREANALAMQKWNLSLDEGRHLTLNLLSKIELRNGEVK 705
| | | | |
RESULT 4
Q9H5J2 PRELIMINARY; PRT; 540 AA.
AC Q9H5J2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ23389.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027042; BAB15634.1; -
SQ SEQUENCE 540 AA; 60158 MW; 8B040AFB87F17A27 CRC64;
Query Match 45.8%; Score 2847; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 3.3e-134;
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 652 MPYADHYSTFSPRDMNSSPYQPPQYPGPVPPVPSGMYAPYDSSRRIRWRRPMYQRDDI 711
| | | | |
Db 1 MPYADHYSTFSPRDMNSSPYQPPQYPGPVPPVPSGMYAPYDSSRRIRWRRPMYQRDDI 60
| | | | |
QY 712 IRSNLSLPPMDVMHSSVYQTSRLRYSNLSLDGYYSVACQPPSPRITVPLPREPCGHLKTS 771
| | | | |
Db 61 IRSNLSLPPMDVMHSSVYQTSRLRYSNLSLDGYYSVACQPPSPRITVPLPREPCGHLKTS 120
| | | | |
QY 772 EQQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPPTPPSPFSDVDFRADFSSESVGTKEED 831
| | | | |
Db 121 EQQIRRKPDQWQAQYHTQKAPLVSTLPVATQSPPTPPSPFSDVDFRADFSSESVGTKEED 180
| | | | |
QY 832 HLSHYSPWSCGTIGSCINAIIDSEPKOVIANNAVLMDLDSGDVKKRRVHLFETQRTKEED 891
| | | | |
Db 181 HLSHYSPWSCGTIGSCINAIIDSEPKOVIANNAVLMDLDSGDVKKRRVHLFETQRTKEED 240
| | | | |
QY 892 PIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLISVSDYVPYVNAVDSRW 951
| | | | |
Db 241 PIIPFSDGPIISKWGAISRSTGYHTTDPVOATASQGSATKPLISVSDYVPYVNAVDSRW 300
| | | | |
QY 952 SSVGNEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQQAQKSNLLQREANALAM 1011
| | | | |
Db 301 SSVGNEATSSAHYVERDRFIVTDLSGHRKHSSTGDLSSLELQQAQKSNLLQREANALAM 360
| | | | |
QY 1012 QKWNLSLDEGRHLTLNLLSKIELRNGELQSDYTEDATDTKPDRIEELSALDTPDQ 1071
| | | | |
Db 361 QKWNLSLDEGRHLTLNLLSKIELRNGELQSDYTEDATDTKPDRIEELSALDTPDQ 420
| | | | |
QY 1072 QSEPIEELDIQLGISNDQLLNGWAVNGHPVQHQKPPKQKQSGEDHVLIEBQK 1131
| | | | |

```
Db 421 QSEPIEELDIQLGISSQNDQLNGMAVENGHPVQHQKEPPKQKQSLGDHVLBEQK 480
Qy 1132 TLIPVTSQSPQLPVSISNASCLPITTSVAGNLILKTHVMSDQNDFLKPVANGKYN 1191
Db 481 TLIPVTSQSPQLPVSISNASCLPITTSVAGNLILKTHVMSDQNDFLKPVANGKYN 540

RESULT 5
Q69231
ID Q69231 PRELIMINARY; PRT; 1136 AA.
AC Q69231;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MKIAA2025 protein (Fragment).
GN Name=mkIAA2025;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Seino S., Nishimura M., Kaisho T., Hoshino K., Kitamura H.,
RA Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene:
RT IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous
RT cDNAs Identified by Screening of Terminal Sequences of cDNA Clones
RT Randomly Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:205-218(2004).
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AK173335; BAD32613.1; -
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:001567; P:protein ubiquitination; IEA.
DR InterPro; IPR00571; Znf_CCH.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; PF00642; zf-CCCH; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf_C3H1; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00899; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
FT NON TER 1
SQ SEQUENCE 1136 AA; 125988 MW; 4A3A2BA6F13C0D51 CRC64;

Query Match 43.2%; Score 2685; DB 2; Length 1136;
Best Local Similarity 52.2%; Pred. No. 1.2e-125;
Matches 615; Conservative 135; Mismatches 295; Indels 134; Gaps 36;

Qy 1 MPVQAQWTEFLSCPICYNEFDNRVHKPISLGGSHYVCKTCLNKLHKRACPPDQTINTD 60
Db 7 MPVQAQWTEFLSCPICTQTFDTIRKPSLGGCHTVCKNKLHKRACPPDQTINTD 66
Qy 61 IDVLPVNFALLQVLGAQVDPHQSIKL-SNLGENKHYEVAKCKVEDIALYLKPLSGGKGYA 119
Db 67 IELLPVNSALLQVLGAQVPEQQPITLCGVEDTKHYEEAKCKVEELALYLKPLSSARGVG 126
Qy 120 --SLNQSALSRPMQRKLVTLVNCQLVEERGRVAMRAARSLGERTVTTEILQHONPQOLS 177
Db 127 LNSTTQSVLSRPMQRKLVTLVHCQLVEEGRIRAMRAARSLGERTVTTEILQHONPQOLS 186
Qy 178 ANLWAAVRAGCCOFLGPAQMEERALKVLVLALEDGSLSRKVLVFLVQRLPRFPQASKT 237
Db 187 SNLWAAVRAGCCOFLGPAQMEERALKVLVLALEDGSLSRKVLVFLVQRLPRFPQASKT 246
Qy 238 SIGHVQLLYRASCFFVKTKRDEDSLSMLQKEEFSYEALRRREHDAQIVHIAMEAGLRISP 297
Db 247 SIGHVQLLYRASCFFVKTKRDEDSLSMLQKEEFTYEALRRHDSQIVQIAMEAGLRAP 306
```

RESULT 6

```
Q81VE6
ID Q81VE6 PRELIMINARY; PRT; 1109 AA.
AC Q81VE6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KIAA2025 protein (Fragment).
GN Name=KIAA2025;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```

NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RA TISSUE=Brain;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB095945; BAC23121.1; -.
DR HSP; P22893; 1M90.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00642; zf-CCH; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 1109 AA; 122994 MW; 1D0D6A5F3E1497EC CRC64;

Query Match 41.5%; Score 2577.5; DB 2; Length 1109;
Best Local Similarity 51.3%; Pred. No. 2.7e-120;
Matches 595; Conservative 135; Mismatches 290; Indels 139; Gaps 37;

QY 25 VHKPISLGSCHTVCCKLKLKHKACPFQDTAINTDIDVLVFNALLQLVGAQVPHQSI 84
DB 1 IRKPISLGGCHTVCCKLKLKHKACPFQDTINTDIELLPNSALLQLVGAQVPEQPI 60

QY 85 KL-SNLGENKHYEAKKVEDLALYLKPLSGGKGA--SLNQSALSRRMQRKLVTLVNCQ 141
DB 61 TLCSGVEDTKHYEAKKVEELALYLKPLSSARGVGLNLTQSVLSRPMQRKLVTLVHCQ 120

QY 142 LVEEGRVTPAARAASLGRRTVTELLQHNQFQQLSANLWAAVARGCOFLGPAHQEAL 201
DB 121 LVEEGRTPAARAASLGRRTVTELLQHNQFQQLSSNLWAAVARGCOFLGPAHQEAL 180

QY 202 KLVLLALEDGSALSRKLVLFVVORLEPRFPQASKTSIGHVQLLYRASCFFVKTRDEDS 261
DB 181 KLVLLALEDGSALSRKLVLFVVORLEPRFPQASKTSIGHVQLLYRASCFFVKTRDEDS 240

QY 262 SLMOLEKEERSYALRREHDAQIVHTAMEAGLRISPEOWSSLLYGLDAKSHMQSIIDKL 321
DB 241 SLMOLEKEERTYALRREHDSQIVQAMEAGLRAPDQWSSLLYGDQSKSHMQSIIDKL 300

QY 322 QSPESPAKSVQELTIVLQRTGDPANLRLRPHLELLANDPNPDVAPSTWEOLENNMVAV 381
DB 301 QTPASPAQSVQELTIALQRTGDPANLRLRPHLELLANDPSPDAPPPTWEOLENGLVAV 360

QY 382 KTVVHGLVDFIQNSRKGHETPOPNKYKTSMCRDLRQCGCRGTNCTFAHSGEELE 441
DB 361 RTVVHGLVDYIQNHSGKGGADQQQPPQHSKITYMKRDMKQRCGRGCPKASCCTFAHSGEELE 420

QY 442 KYRLRNKKINATVTRPPLNKKVGNVNTVTTAGNVISVIGSTETGKIYPS-TNGISNAE 500
DB 421 KFKKNNKRL---VPRPLSASLQQLNEVGLPSAAILPDGSAVDLPSPKPPALPNGIVSTG 477

QY 501 NNSVQLISSTSTLRALETVKVKGKVGANGQNAAGPSADSVENKIGSPPTKPPVSNVAA 560
DB 478 NTVTQIPIRGTPD---SYDSSLKPGKI-----DHLSSAPGSPPLLESVPKS 522

QY 561 TSAGPSNVGTSLNVPQKSSPFLTRVPV-----YPPHSENIQYQDPR-TQI 606
DB 523 ISALFVNP---HSIPRGPADLPMPVTKPLQMPVRGSQLYPAQQTQD-YQDDPKGAAP 577

QY 607 PFEVPOYPTGYPPPTTPAGVAPCVPRFVRSNNVPESSLPASPADYADHYSTFSPDR 666
DB 578 PEPAPYQGMYYTPP-----QCVSRFVR-----PPPSAEPAPYLDHPPYL-QER 625

QY 667 MNSSPYQPPPPQYGPVPPVPSGMYAPVYDSRIWRPMPYQRDITRSNLSLPPMDVMHGS 726
DB 626 VVNSQYGTQPOQ-----YPII-----YPSHYDGRVYAPSVYTRIEIPRESPI-PIEIPPA 676

QY 727 V--YQTSLRERYNSLDGYYSVACOP-----PS-----EPRTTVPLPREPGHLKTSCEQIR 776
DB 677 VPSYVPESERYQQIETSYYPVAPHPTQIRPSYLRPPYSRLPPPPQP-----HPSLDELHR 732

QY 777 RKPDOMAQYHTQKAPLVSSTLPVATQSPTPSPPLFSVDPRADFSSESVCSTKFEEDHLSHY 836
DB 733 RKREIMAQLERK---VISPPFFA-PSPTLP-PTPHPEFLDEDLKVAG-KYKGNDSQY 786

QY 837 SPWSCGTIGSCINAIIDSEPKDVIANSNAVLMDLSDGDVKKRRVHLFETQRTKE--EDPII 894
DB 787 SPWSCDTIGSYIGTKDAKPKDVAAGSVEMNVESKGM--RQRLDLQRAAETSDDDLI 844

QY 895 PFDGPIISKWCAISRRTGHTTDPVQATASQGSATKPSVSVYVFNVAVDNRW--S 952
DB 845 PFGDRPTVSRFGAISRTSKTIYQAGAPMQAMAPQAGPTKSNISIDSPY--GTHGGWGAS 902

QY 953 SYG--NEATSAHYVERDRFVTDLSGHRK--HSSTGDLALLSLELQO----- 994
DB 903 PYSHPQNTPSQGHFSEERISMSVASHGKPLPSAREQLRLELQQLNHOISQQTQLRGL 962

QY 995 -AKNSNLLQREANALAMQO-----KWNLSDEGRHLTLNL--LSKEIELRNGELQSD 1043
DB 963 EAVSNRLVLQREANTLACSQSPPPPPKMPCGMISSEQLSLELHQVREIGKRTRELSME 1022

QY 1044 YTEDATDTKPRDIELELSALDTDE--PDGQSEPIEEILDIOLGISSQNDQLNGAVEN 1101
DB 1023 -NQCSLDMK-----SKLNTSKQAEAGQPEPQNKV-----PAEDLTFTSDVPN 1064

QY 1102 GHPVQOHOKEPPKQKQSL 1120
DB 1065 GSALTQENISLLSNKTS 1083

RESULT 7
Q6NUC6 PRELIMINARY; PRT; 1114 AA.
AC Q6NUC6;
ID Q6NUC6;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE MGC81061 protein.
GN Name=MGC81061;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hoig F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Hellon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
```

RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Klein S., Strausberg R.,
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; BC068669; AAH66669.1; -
DR GO; GO:0000153; Cubiquitin ligase complex; IEA.
DR GO; GO:0003676; F-nucleic acid binding; IEA.
DR GO; GO:0004842; F-ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F-zinc ion binding; IEA.
DR GO; GO:0016567; P-protein ubiquitination; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR Pfam; PF00642; zf-CCHC; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf_C3H1; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 1114 AA; 123576 MW; E836C7514BF2A3B7 CRC64;

Query Match 39.9%; Score 2477; DB 2; Length 1114;
Best Local Similarity 47.3%; Pred. No. 2.8e-115;
Matches 592; Conservative 140; Mismatches 310; Indels 210; Gaps 40;

QY 1 MPVQAQMTFELSCPIYNEFDENVKPISLCSHTVCKTLNKLKHKACPFDPQALNTD 60
DB 1 MPVQAQMTFELSCPICTQTFDTETIKPISLGGHTVCKWCLNKLKHKACPFDPQITNTD 60

QY 61 IDVLVNFALLQVGAQVDPHQSIKL---SNLGENKHVEAKVCVEDLALYLPKLSGGKG 117
DB 61 IELLVPNSALLQVGAQVPEQQOITLCCGCGAEDTKHYEARKCVELALYLPKLSIARG 120

QY 118 VA--SLNQSALGRPMOKLVLVNCOLVEEGRVRAARAARSIGERTVTTELILQHNPQQ 175
DB 121 VGENSTQSVLRPMQKLVTLVHQLVVEEGRIRARAARSIGERTVTTELILQHNPQQ 180

QY 176 LSNLWAAVRAQCQFLGPAMQEEALKVLLALEDGSALSRKVLVLFVVQRLEPRPQAS 235
DB 181 LSNLWAAVRAQCQFLGPAMQEEALKVLLALEDGSALSRKVLVLFVVQRLEPRPQAS 240

QY 236 KTSIGHVQVLLYRASCFTYKREDESSLMOLEKFRSVEALREHDAQIVHTAMEAGLRI 295
DB 241 KTSIGHVQVLLYRASCFTYKREDESSLMOLEKFRSVEALREHDSQIVQVAMEAGLRI 300

QY 296 SPEOWSSLLYGLAHKSHMQSIIDKLOSPEFSAKSVOELTI VLORTGDPANLRLRPHLE 355
DB 301 AFDQWSSLLYGOQSHKSHMQSIIDKLOT PASFAQSVOELTIALORTGDPANLRLRPHLE 360

QY 356 LLANIDPNPDVSPTEQLENAMAVKTVVHGLVDFIQNYSRKGHETPQPNSKYKTSM 415
DB 361 LLANIDPSPDAPPTTEQQLDGLVAVKTVVHGLVDVIQNHKKGTDDQQQPQHSKYKTM 420

QY 416 CRDLROGCGPRGTNCTFAHSELEKYLRLNKNINATVTPPL-----LNKVGNNVT 469
DB 421 CRDMKQRCGPRGASCTFAHSELEKFLRMNKL---VPRPLSASLQQLNEVGL----- 473

QY 470 TTAGNVISVSTETTGKIVP--STNGISNA---ENSVQSILSRSTDTSLRALETVKV 524
DB 474 -----PIGAPDEGMDLPRKSGSLPNGIVPGSSVTQLISRSTDS---GFESVLKP 521

QY 525 GKVGANGQNAAGSADSVTENKIGSPKTPVSNV-----AATSGAPNVGTGLNSVPQ 577
DB 522 VKLDHLSSAPGSPPE-----LIDSVPKTSMSALPVNAHPAASRDLPPLFVSKQIQWVPR 576

RESULT 8

Q86ST6
ID Q86ST6 PRELIMINARY; PRT; 506 AA.
AC Q86ST6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-WAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Casavant T.L., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.L., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; BC044642; AAH44642.1; -;
DR HSP; P22893; IM90.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR000571; Znf_CCH.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR Pfam; PF00642; zf-CCCH; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00356; Znf_C3H1; 1; UNKNOWN_1.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 506 AA; 56853 MW; F2915DA0B8D0AC9E CRC64;

Query Match 36.9%; Score 2293.5; DB 2; Length 506;
Best Local Similarity 95.5%; Pred. No. 1.4e-106;
Matches 448; Conservative 3; Mismatches 13; Indels 5; Gaps 1;
QY 1 MPVQAAQWTEFLSCPICYNEFDENVHKPISLGCSTVCKTCLNKLRRKACPDQTAINTD 60
DB 1 MPVQAAQWTEFLSCPICYNEFDENVHKPISLGCSTVCKTCLNKLRRKACPDQTAINTD 60
QY 61 IDVLVPNFALLQVGAQPDHQSILSNLGENKHVEAKKVEDLALYLKPLSGGKGVAS 120
DB 61 IDVLVPNFALLQVGAQPDHQSILSNLGENKHVEAKKVEDLALYLKPLSGGKGVAS 120
QY 121 LNQSALSRLPMQKLVTLNQCQLVEEGRVRAARAISLGERVTTELILQHNPQOLSANL 180
DB 121 LNQSALSRLPMQKLVTLNQCQLVEEGRVRAARAISLGERVTTELILQHNPQOLSANL 180
QY 181 WAAVRAGCQFLGPMQOEALKVLVLALEDGALSRLKVLVFWORLEPRFPQAKTSIG 240
DB 181 WAAVRAGCQFLGPMQOEALKVLVLALEDGALSRLKVLVFWORLEPRFPQAKTSIG 240
QY 241 HVVQLLYRASCFTVKTKRDESSLMQLKEFRSYEARLREHDAQIVHIAEAGLRISPEQW 300
DB 241 HVVQLLYRASCFTVKTKRDESSLMQLKEFRSYEARLREHDAQIVHIAEAGLRISPEQW 300
QY 301 SLLLYGDLAHKSHMQSIIDKQSPESFAKSQVELTIVLQRTGDPANLRLRPHLELLANI 360
DB 301 SLLLYGDLAHKSHMQSIIDKQSPESFAKSQVELTIVLQRTGDPANLRLRPHLELLANI 360
QY 361 DNPDAVSTPWELENAMAVKTVVHGLVDFDTQNSRGKHETPQOPNSKYKTSKCRDLR 420
DB 361 DNPDAVSTPWELENAMAVKTVVHGLVDFDTQNSRGKHETPQOPNSKYKTSKCRDLR 420
QY 421 QQGGCPRGNTCTFAHSELEKYLRLNKKINATVTRTPFLNKKVGNNTV 469
DB 421 QQGGCPRGNTCTFAHSELEKYLRLNKKINATVTRTPFLNKKVGNNTV 469

RESULT 9
Q9HBD1

ID Q9HBD1 PRELIMINARY; PRT; 419 AA.

AC Q9HBD1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Membrane-associated nucleic acid binding protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20517928; PubMed=10938276; DOI=10.1074/jbc.M004461200;
RA Siebs D.C., Vedder C.T., Merkens L.S., Tanaka T., Fred A.C.,
RA McCoy S.L., Heinrich M.C., Diefenbach M.E., Bennett R.M.,
RA Hefeneider S.H.;
RT "A human gene coding for a membrane-associated nucleic acid-binding
RT protein.";
RL J. Biol. Chem. 275:33655-33662(2000).
DR EMBL; AF255304; AAG00433.1; -;
DR GO; GO:0009986; C:cell surface; IEA.
DR GO; GO:0005624; C:membrane fraction; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
FT NON_TER
SQ SEQUENCE 419 AA; 46285 MW; 2C525A12C77F56DC CRC64;
Query Match 34.8%; Score 2165; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 2.7e-100;
Matches 419; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 773 EQIRKPKDQWQAQYHTQKAPLVSSSTLPVATQPTPSPPLFSVDPRADFSVSQGTKEEDH 832
DB 1 EQIRKPKDQWQAQYHTQKAPLVSSSTLPVATQPTPSPPLFSVDPRADFSVSQGTKEEDH 60
QY 833 LSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVLMDLSDGVKRRVHLFETQRTKEEDP 892
DB 61 LSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVLMDLSDGVKRRVHLFETQRTKEEDP 120
QY 893 IIPESDGP1ISKWGAISRSTGYHTTDPVQATASQGSATKPIVSVDYVYPYNAVDSRWS 952
DB 121 IIPESDGP1ISKWGAISRSTGYHTTDPVQATASQGSATKPIVSVDYVYPYNAVDSRWS 180
QY 953 SYGNEATSAHYVERDRFIVTDLSGHRKHSSTGDLISLELOQAKSNLLQREANALAQ 1012
DB 181 SYGNEATSAHYVERDRFIVTDLSGHRKHSSTGDLISLELOQAKSNLLQREANALAQ 240
QY 1013 QKWNLSDEGRHLTLNLSKEIERNLNGELQSDYTEDATDTKPRDIELELSALDTDEPDQ 1072
DB 241 QKWNLSDEGRHLTLNLSKEIERNLNGELQSDYTEDATDTKPRDIELELSALDTDEPDQ 300
QY 1073 SEPIBEILDIOIGISSQNDQLNGMAVENGHVPVQHQKEPPKQKQSLGEDHVLLEQKT 1132
DB 301 SEPIBEILDIOIGISSQNDQLNGMAVENGHVPVQHQKEPPKQKQSLGEDHVLLEQKT 360
QY 1133 ILPVTSCSQPLPVSISNASCLPITTSVAGNLIILKTHVMSDKNDFLKPVANGKQVNS 1191
DB 361 ILPVTSCSQPLPVSISNASCLPITTSVAGNLIILKTHVMSDKNDFLKPVANGKQVNS 419
RESULT 10
Q9VV48
ID Q9VV48 PRELIMINARY; PRT; 819 AA.
AC Q9VV48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG16807-PA.
GN ORFNames=CG16807;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasse K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupek M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodagat, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou X.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RT Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=2426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence";
RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=2426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective";
RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=2426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review";

RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; AE003527; AAF49474.1; -.
DR FlyBase; FBgn0036621; CG16807.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF00097; Zf-C3HC4_1.
DR Pfam; PF00642; Zf-CCCH; 1.
DR SMART; SM00184; Ring; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 819 AA; 90376 MW; 33D429076EDE7E2E CRC64;
Query Match 24.7%; Score 1538; DB 2; Length 819;
Best Local Similarity 40.5%; Pred. No. 1.5e-68;
Matches 386; Conservative 98; Mismatches 226; Indels 242; Gaps 29;
QY 1 MPVQAQWTEFLSCPICYNFEEDNVHVKPISLGCSHTCKLNLKLRKACPFDDTAINTD 60
DB 1 MPQAPSMTDFELPCICCNFAASQRCVPVSGCGHTICKLCTLYNQCFDQVIVSD 60
QY 61 IDLVFNFALLQLV-----GAQVPDQHSIKLSNLGENKHYEYAKVCVEDLALYLKPL 112
DB 61 IDNLPINHALQLVKDSELLELAPPPSVQKLEPHL--KCYQLGRCIEELALHLKSF 117
QY 113 SGKGVSINOSALSRRPMOKLVTLVNCQVVEEGRVRAARSLGERTVTELLIQHON 172
DB 118 LNLNG-----NGNLLTRPMLRKLVTLVNCQVVEEGRVRAARSLGERTVTELLIQHON 173
QY 173 POOLSANLWAAVRARGCOFLGPAMQEEALKVLVLALEDGSSALSRKVLVLFVVRLEPRFP 232
DB 174 POOLSSNLWAAVRTRGCOFLGPAMQEEVLKVLVLALEDGSSALSRKVLVLFVVRLEPRFP 233
QY 233 QAKSTSIHVQVQLLYRASCVKYTKRDESSLMQLKEEFSYVEALRREHDAQIVHIAMEAG 292
DB 234 QAKSTSIHVQVQLLYRASCVKYTKRDESSLMQLKEEFSYVEALRREHDAQIVHIATEAG 293
QY 293 LRISPEQWSSLLYGDLAHSHMQSIIDKLQSPESFAKSVQELTIVLQRTGDPANLRLRP 352
DB 294 LRIAPEQWSSLLYGDVMEKSHMQSIIDKLQTPSSFAQSVQELTIVLQRTSDPAKLSLHH 353
QY 353 HLELANIDENPDVAVSPVTEOLENAMAVKTVHGLVDFIQNYS-RK-----GHETPQ 404
DB 354 HUKYLANIDPCAE-VAP-WSVLAEALDAVRHSVGLVNFQLHGHVGRKQDGSIGSGSGGT 411
QY 405 POPSNKYKTSIMCRDLRQCGGPRGTNCTFAHSGEELEKYLRLNKKINATVTRTPFLNKKVG 464
DB 412 TNSNPYKISLCRDLVRRVRCPRGSSCTFAHSGEEVERARNR----- 455
QY 455 VNNVTVTAGNVISVIGSTETTKVLPSTNGISNAENSVSQISLRSTDLRALETVKV 524
DB 456 ----- 455
QY 525 GKVGVANGQNAAGPSADSVTENKIGSPKPTPSNVAAATSAGPSNVGTSLNSVP--OKSSSP 582
DB 456 -----GKMKTPLA-----LOGPPAVGVGAIKKPLGEGEGPP 487
QY 583 LTRVFPVPPHSENIQYFODPRTQIPFVFPVQYPTGYPPTPTVPAGVAPCVPRFRVSNV 642
DB 583 LTRVFPVPPHSENIQYFODPRTQIPFVFPVQYPTGYPPTPTVPAGVAPCVPRFRVSNV 642

```
Db 488 LGNMPMLPMPSP-MHYMGSPR-----GYLDP-----SLGLSP-----GGGL 522
QY 643 PESSLPPAS---MP--YADHYSTF---SPDRMNSSPYQPPPPQYGPVPPVPSGMYAPV 694
Db 523 PFSHISPIRLVPSYDRSFGGGR-----IPSPREYQANPVAP-----566
QY 695 YDSRIWRPMPYQORDIIRSNLPPM---DVMHSSVYQTSRLRYSNLDGYYSVACQPP- 750
Db 567 --TORANPNFNSVSNLHKGYMLPASGGDVFLA-----NPMQAYLAQQQHP 614
QY 751 -----SEPTTVPPLPREPCGHLKTSCEQ-IRKPDQWQYHTQKAPLVS-STLPVATQS 803
Db 615 QHPQQQQPSSKPNRPLSILPATADTSFFKCPNPSVSIIDLRVPEVNVDAVPLFRS 674
QY 804 PTPPSPLFSDPRADPSESVGTFKFEEDHLSHYSPWSCGTIGSCINAIIDSEPKDVIANSN 863
Db 675 -----NNNNNNNNNNNNNNHSSLLFPN-NTGKDSANFVRS---DSLDDDD 720
QY 864 AVLMDLDSGDVKKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRRTG 915
Db 721 ASTFDVPTGSSMLSIY-----GPICPK-----SSTTG 747

RESULT 11
Q7PTD3
ID Q7PTD3 PRELIMINARY; PRT; 398 AA.
AC Q7PTD3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000011168 (Fragment).
GN Name=ENSANGG0000008679;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PEST;
RC Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: AAB01008807; EAA04070.2; -.
DR GO; GO:0003676; F: nucleic acid binding; IEA.
DR InterPro; IPR000571; Znf_CCHC.
DR Pfam; PF00642; zf-CCHC; 1.
FT NON_TER 1
FT NON_TER 398
SQ SEQUENCE 398 AA; 43739 MW; F1BCAFD08815A328 CRC64;

Query Match 20.7%; Score 1289.5; DB 2; Length 398;
Best Local Similarity 67.2%; Pred. No. 1.4e-56;
Matches 268; Conservative 38; Mismatches 68; Indels 25; Gaps 5;

QY 82 QSLKLNGLGNKHYVAKKVEDLALYLP-----LSGGKG-----VASLNOSALS 127
Db 2 QSLSPENL---QYKSAKGCIIEALALYLKPCPSAGSGSVAGSGNGVGGSGSPSGASLLS 58

QY 128 RPNQRLVTLVNCQVLEEGRVRAARAARSLGRTVTTELQHNQPNQOLSANLWAAVRAV 187
Db 59 RQNRQLVTLVNCQLIEDGRARSLRAARSLGRTVTWELIHHQNLQOLSTNLWAAVRAV 118

QY 188 GCQFLGPMQEEALKVLVLALEDGSAISRKVLVFLVVQRLEPPQASKTSIGHVQVLLY 247
Db 119 GCQFLGPMQEEVLKVLVLALEDGSAISRKVLVFLVVQRLEPPQASKTSIGHVQVLLY 178

QY 248 RASCFKVKTKDESSLMQLEERSVEALRREDAQIVHTAMEAGLRISPEQWSSLLYGD 307
Db 179 RASCFKVKTKREGSSLMQLEERTVEALRREDAQIVQIATEAGLRAPDQWSSLLYGD 238

QY 308 LAHKSQMSIIDKLSQSPESPAKSVQELTIVLQRTGDPANLRLRPHLELLANLDPNDAV 367
```

```
Db 239 TAHKSHMQSINDKLTQTPQSFQSVQELIILQRTQDPAKSLRVLQKHLAIDNSEN 298
QY 368 SPTVEQLENAVAVKTVVHGLVDFQNY-SRKGHETPQPPNPKYKTSKCRDLRQGGCP 426
Db 299 VPSMAECAALQAVKRVVAGLVDFVQHGNRKLQEAAGLAHNRNYKLSICRDLNNGTGP 358
QY 427 RGTNCTFAHSGEELKYLKRNKK-----INATVTRTP 458
Db 359 RGPNCPTFAHSGEELKYLKRNKK-----INATVTRTP 397

RESULT 12
Q9NNW9
ID Q9NNW9 PRELIMINARY; PRT; 198 AA.
AC Q9NNW9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ20713.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
RA Shibahara T., Tanaka T., Nakamura Y., Isegaki T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK00720; BAA91340.1; -.
SQ SEQUENCE 198 AA; 22234 MW; E7778660861F1960 CRC64;

Query Match 14.3%; Score 889; DB 2; Length 198;
Best Local Similarity 98.9%; Pred. No. 5.3e-37;
Matches 173; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 867 MDLSDGDVKKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRRTGYHTTDPVQATA 926
Db 1 MDLSDGDVKKRRVHLFETQRTKEEDPIIPFSDGPIISKWGAISRRTGYHTTDPVQATA 60

QY 927 SQGSATKPIVSVDYPPYNAVDSRWSYSGNEATSSAHYVERDRFVTDLSGHRKHSSTGD 986
Db 61 SQGSATKPIVSVDYPPYNAVDSRWSYSGNEATSSAHYVERDRFVTDLSGHRKHSSTGD 120

QY 987 LLSLELQQAQKNSLLQREANALAMQKQWNSLDEGRHLTLNLSKEIELRNGELQ 1041
Db 121 LLSLELQQAQKNSLLQREANALAMQKQWNSLDEGRHLTLNLSKEIELRNGEVK 175

RESULT 13
O45962
ID O45962 PRELIMINARY; PRT; 1014 AA.
AC O45962; Q93898;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein M142.6.
GN ORFNames=M142.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=Bristol N2;
RA Steward C.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurray A.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 273428; CAB16481.2; JOINED.
DR EMBL; 299276; CAB16481.2; -.
DR EMBL; 273428; CAA37810.2; -.
DR EMBL; 299276; CAA37810.2; JOINED.
DR F1; T23764; T23764.
DR IntAct; O45962; -.
DR WormBase; WBGene00010923; M142.6.
DR WormPep; M142.6; CE34061.
DR GO; GO:000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0014567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; ZnF_ring.
DR PROSITE; PS00518; ZF_RING_1; UNKNOWN_1.
DR PROSITE; PS00589; ZF_RING_2; 1.
KW Hypothetical protein.
SQ
SEQUENCE 1014 AA; 112311 MW; 7226503B2DEB4008 CRC64;

Query Match 13.6%; Score 843; DB 2; Length 1014;
Best Local Similarity 30.6%; Pred. No. 1.1e-33;
Matches 238; Conservative 99; Mismatches 278; Indels 162; Gaps 23;

QY 4 QAAQTEFLSCPTCYNEFDENVHKKIPISLGCSTHVCKTCLNKLHR--KACPFDTAINTDI 61
DB 6 QGGQWQEVLCSSICNRHFE-TFLPVSLICGHVICKKAEKPNQKPCPHDQWKTHSP 64

QY 62 DVLVFNALLQLVGAQVDPHQSIKLSNLGENKHGKHYEYAKKCVEDLALYKPLSGGKGVASL 121
DB 65 SEYPNNVALLSVI---FPRKQCWTLGSAVSEAEKRVQDLSIQ-IAKFFREADSERG-GTV 119

QY 122 NSGALSRRPQKLVTLVNCQVVEEGRVRAMRAARSLGRTVTTELILQHONQQQLSANLW 181
DB 120 SSREISRTLQKRVALLCYQREVQDGLKTLKMCGRGISERVMEIILSIQSNTHVSSQLW 179

QY 182 AAVRARGCQFLGPMQDEEALKLVLALDEGSSALSRKVLVFLVVORLEPRFPQASKTSGH 241
DB 180 SAVRARGCQFLGPMQDDVRLILMTLETGECARKNLVYVVTLASYPQVKTCVGH 239

QY 242 VVQLLYRASCYKVKTRDESSLMQKEEPRSYEARREHDAQIVHIAMEAGLRISPEQWS 301
DB 240 VVQLLYRASCYVNLKRGESSLMQKEEPRSYEARREHDSQIVQIAPESGLRIGPDQWS 299

QY 302 SLLYGLDAKSHMQSIIDKQSPESPAKSVQELTIVLQRTGDPAN-----LNRLRP 352
DB 300 ALLYADQSHRSHMQSIIDKQSKNSYQQGVVEELRAL---AGSQTSLMLVPAYRYFLTQVIP 356

QY 353 HLELLANIDPNDAVSPTEQLENAMAVKTVVHGLVDFIONV---SRGHETPQPQNS 409
DB 357 CLEFFAGIEHEDTSMNIGDALHQIRILK--LHCSQDDLRKMPKEERGVLQAEVPGG 414

QY 410 KYKTSNCRDLRQGGCPRGNTCTTFAHSQBELEKYLRLNKKINATVTRTFFLLNKVGNNTV 469
DB 415 -----MGGGPGSG-----GAERAGRI----- 430

QY 470 TTAGNVISVIGSTETTKIVSTNGISNAENSVSQILSRSTDSTLRALETVKVKGKGA 529
DB 431 ----GGLHPLYSQIDETGRSISRTNPKNSHNSPQ----- 461

QY 530 NQNAAGPSADSVTENKISGPPKTPVSNVAATSAGPSNVGTLSNVQKSSPLTRVPVY 589
DB 462 -----TPPKQPKRYQMGIPIPNRMG----- 497

QY 590 PPHSENIQFQDPRTOIPFEVPOYPTGYPPPTVPAGVPCVPRFVRNNVPESLPP 649
DB 498 PPP-----QPFNS--QHLPRFRGRGRGAPPPPPPP-----MPMLIGVD-----MPG 539
```

```
650 ASMPYADHYSTFSPDRMNSSP-----YQPPPPQVGVPPVPSPGMYAPVYDSRRIRWPP 704
540 APMQATEVLT-ADQMVGNGTPQRVVMQSPHLPGGVVMVPPQQMQVPPQSPMTVPVGGP 598
705 MYQRDDIIRSNLPPMDVMHSSV-YQTSRLRERYNSLDGYYSV---ACOPPSPPRTTV 757
599 M-----GPMGPMTFPIPVQPPNTMTATSPGTSVIVYPAASPPGQPPHTI 643

RESULT 14
Q96F27
ID Q96F27 PRELIMINARY; PRT; 177 AA.
AC Q96F27;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.J., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska J., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011688; AAH1688.2; -.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;

Query Match 12.6%; Score 780; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 1.3e-31;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LNKLRKACPFDTAINTDIDLVPVNFALLQVGAQVDPHQSIKLSNLGENKHGKHYEYAKK 101
DB 1 LNKLRKACPFDTAINTDIDLVPVNFALLQVGAQVDPHQSIKLSNLGENKHGKHYEYAKK 60

QY 102 VEDLALYKPLSGGKGVASLNSQALSRRPMQKRLVTLVNCQVVEEGRVRAMRAARSLGR 161
DB 61 VEDLALYKPLSGGKGVASLNSQALSRRPMQKRLVTLVNCQVVEEGRVRAMRAARSLGR 120

QY 162 TVTELIQHONQQQLSANLWAAVRARGCQFLGP 194
DB 121 TVTELIQHONQQQLSANLWAAVRARGCQFLGP 153
```


This Page Blank (uspt)